PATENTTI- JA REKISTERIHALLITUS
NATIONAL BOARD OF PATENTS AND REGISTRATION

Helsinki 9.2.2004

E T U O I K E U S T O D I S T U S
P R I O R I T Y   D O C U M E N T

Hakija Applicant	**1. Wahlroos, Tony, Turku, FI** **2. Atabekov, Joseph, Moscow, RU** **3. Dorokhov, Yurii, Moscow, Ru** **4. Susi, Petri, Turku, FI** **5. Mäkelä, Mauri, Turku, FI** **6. Korpela, Timo, Turku, FI**
Patenttihakemus nro Patent application no	**20030315**
Tekemispäivä Filing date	**28.02.2003**
Kansainvälinen luokka International class	**C12N**
Keksinnön nimitys Title of invention	

"Methods and constructs for increasing the content of selected amino acids in seeds"
(Menetelmiä ja konstruktioita valikoitujen aminohappojen pitoisuuksien lisäämiseksi siemenissä)

Hakemus on hakemusdiaariin **20.05.2003** tehdyn merkinnän mukaan siirtynyt **Boreal Plant Breeding Ltd, Jokioinen.**

The application has according to an entry made in the register of patent applications on **20.05.2003** been assigned to **Boreal Plant Breeding Ltd, Jokioinen.**

Täten todistetaan, että oheiset asiakirjat ovat tarkkoja jäljennöksiä Patentti- ja rekisterihallitukselle alkuaan annetuista selityksestä, patenttivaatimuksista, tiivistelmästä ja piirustuksista.

This is to certify that the annexed documents are true copies of the description, claims, abstract and drawings originally filed with the Finnish Patent Office.

Pirjo Kaila
Tutkimussihteeri

Maksu 50 €
Fee 50 EUR

Maksu perustuu kauppa- ja teollisuusministeriön antamaan asetukseen 1027/2001 Patentti- ja rekisterihallituksen maksullisista suoritteista muutoksineen.

The fee is based on the Decree with amendments of the Ministry of Trade and Industry No. 1027/2001 concerning the chargeable services of the National Board of Patents and Registration of Finland.

Osoite: Arkadiankatu 6 A   Puhelin: 09 6939 500   Telefax: 09 6939 5328
P.O.Box 1160   Telephone: + 358 9 6939 500   Telefax: + 358 9 6939 5328
FIN-00101 Helsinki, FINLAND

Methods and constructs for increasing the content of selected amino acids in seeds

Technical Field of the Invention

The present invention relates in general to plant biotechnology. Particularly, the invention is related to methods and constructs for increasing the amino acid content in any plant species and in any plant tissues especially in cell walls and membranes, and particularly in seeds. Compositions obtainable by the method and the use of said amino acid-enriched composition are disclosed.

Background of the Invention

Human beings and livestock require eight essential amino acids in their diets. Diets based predominantly on a single cereal or legume species result in amino acid deficiencies due to nutritional limitation of seed proteins that may have a negative effect(s) on the dietary needs of human beings and animals. For example, the proteins in cereal seeds are deficient in lysine and tryptophan, whereas legume seeds contain proteins deficient in the sulfur-containing amino acids, methionine and cysteine. The use of seed proteins in feed of livestock necessitates that the diet has a prescribed amino acid composition in order to promote the health of animals, efficient growth and good quality of meat and milk. Therefore, it is advantageous to modify existing plant protein resources, in particular, for the composition of essential amino acids in order to be better adapted to the needs of a specified animal.

Efforts have been made to match the composition of vegetable amino acids to the dietary needs of humans and animals, but with limited success. For example, the use of nutritionally superior plant mutants and tissues thereof, is compromised by negative pleiotropic effects. These problems include poor seed germination, slow dry-down, reduced yield, increased microbial and insect susceptibility, and poor milling characteristics.

Genetic engineering provides an alternative means of changing the composition of essential or any amino acid in plants and tissues thereof. In order to elevate methionine or lysine, either biosynthetic pathways have been manipulated, or high-methionine/lysine proteins have been

expressed in transgenic seeds (WO 96/38574; WO 96/01905; WO 95/31554; WO 95/15392; WO 93/19190; EP 485 970, WO 99/40209). Principally, three methods have been applied: 1) the amino acid sequence of endogenous protein(s) was altered; 2) naturally occurring proteins from other plant species were recruited for heterologous expression; and 3) synthetic genes containing high levels of methionine/lysine were expressed.

The method of using high-methionine/lysine proteins from other plants has been disclosed in US 5,633,436, US 5,580,782, WO 94/16078). With the aim of improving the nutritive value of an important grain legume crop, a chimeric gene specifying seed-specific expression of a sulfur-rich, sunflower seed albumin was stably transformed into narrow-leafed lupin *(Lupinus angustifolius* L.). Sunflower seed albumin accounted for 5% of extractable seed protein in a line containing a single tandem insert of the transferred DNA. The transgenic seeds contained less sulfate and more total amino acid sulfur than the non-transgenic parent line; this was associated with a 94% increase in methionine content and a 12% reduction in cysteine content. There was no statistically significant change in other amino acids or in total nitrogen or total sulfur content of seeds.

A combination of the first and second approaches (above) has been described in US 5,850,016. To improve the methionine content of potato tubers, a cDNA clone encoding Brazil nut 2S albumin was mutagenized to increase its methionine content by 2-7 additional methionine residues and transformed into potato plants. Irrespective of the mutation, protein content in leaves was low, ranging from < 0.01%-0.2% of total protein.

The use of synthetic proteins has been disclosed in FR 2,744,134; US 5,559,223; WO 92/14822). To increase the lysine and methionine content in seeds, a synthetic protein based on an α-helical coiled-coil structure containing 31% lysine and 20% methionine (CP3-5) was designed. Driven either by the phaseolin or ß-conglycinin promoter, moderate amounts of the synthetic protein accumulated in seeds harvested from transgenic tobacco plants.

In WO 99/15004 a chimeric construct for modifying the composition of storage organs in plants is described. A gene encoding a sulphur-rich protein is provided with a C-terminal KDEL extension, which enables targeting of the construct to the endoplasmic reticulum and

Golgi apparatus.

The above approaches have certain drawbacks. When large amounts of foreign proteins with no functional role for the host plant are expressed it may result in many secondary problems connected with physiological abnormalities of seeds as noticed in equivalent-type mutants created by traditional breeding methods.

The first objective of the present invention is to provide a means, including recombinant nucleotide sequence constructs, enabling selection of constructs for effective transformation of any desired plant species with targeted expression enabling accumulation of stable protein enriched by any amino acid in any selected plant tissues. When using said constructs the detrimental effects of expressing large amounts of foreign proteins with no functional role for the host plant can be avoided. A second objective of the present invention is to provide a composition comprising a stable amino acid-enriched protein, which has accumulated in selected plant tissue combined with compatible formulation aiding additives. The use of said composition to be suitable as a direct food source from selected plant for humans as well as fodder, especially as animal feed and as a feed supplement, is suggested

Summary of the Invention

The characteristic features of the present invention are as defined in the claims.

A Brief Description of the Drawings

Figure 1 shows eighteen expression plasmids constructed using either of two carrier-proteins (Ole and TMV MP) with one of three transcriptional promoters indicated by arrows (NAP, HYB, 35S CaMV) and internal his or cys-met-enriched regions of varying length (2x, 4x, and 6x). The green fluorescent protein (GFP) gene was fused in-frame to sequences enriched with his-codons. The transcriptional termination sequence is indicated by "t" in a small box.

Figure 2 shows four expression plasmids selected in transient expression assays for transformation of *Brassica campestris* plants.

Figure 3 shows the nucleotide sequence of the his-codon-enriched sequence (A), the nucleotide sequence of the cys-met-codon-enriched sequence (B), the nucleotide sequence of the glycine-codon-enriched sequence (C) and the nucleotide sequence of the lysine-codon enriched sequence (D). The protein translation sequence is shown below the nucleotide sequence. K, lysine; H, histidine; R, arginine, V, valine, G, glycine, L, leucine, C, cystine, S, serine, M, methionine. The asterisk indicates the transcriptional termination codon. Restriction sites are shown in bold and are indicated above the sequence.

Figure 4 shows the construction scheme of plasmids carrying fusions of histidine or cystine-methionine codon-enriched DNA sequences (gray boxes) of different lengths fused to the TMV 30K MP gene (open box "MP").

Figure 5A shows the construction of expression plasmids containing the napin promoter or chimeric promoter that includes the entire napin promoter coupled with the enhancer region of the CaMV 35S promoter.

Figure 5B depicts the cloning of the GFP gene into 35S and napin promoter-based plasmids. Arrows indicate promoters, small box labelled as "t" indicates the transcription termination sequence.

Figure 6 shows the relative GFP expression achieved using the 35S (a) and HYB (b) promoters after particle bombardment of constructs into epidermal cells of *Nicotiana benthamiana*.

Figure 7 depicts the construction of plant expression vectors carrying fusion proteins consisting of the TMV MP, his or cys-met-enriched sequences of different length, and GFP, under the control of the 35S CaMV promoter.

Figure 8 depicts the construction of plant expression vectors carrying fusion proteins consisting of the TMV MP, his or cys-met-enriched sequences of different length, and GFP, under the control of the napin promoter

Figure 9 depicts the construction of plant expression vectors carrying fusion proteins consisting of the TMV MP, his or cys-met-enriched sequences of different length, and GFP, under the control of the hybrid promoter

Figure 10 depicts the construction of plant expression vectors carrying fusion proteins consisting of oleosin, his or cys-met-enriched sequences of different length, and GFP, under the control of the 35S CaMV promoter.

Figure 11 depicts the construction of plant expression vectors carrying fusion proteins consisting of oleosin, his or cys-met-enriched sequences of different length, and GFP, under the control of the napin promoter.

Figure 12 depicts the construction of plant expression vectors carrying fusion proteins consisting of oleosin, his or cys-met-enriched sequences of different length, and GFP, under the control of the hybrid promoter.

Figure 13 shows a Western blot of Nap-MP expressing *Brassica campestris* lines. Lane number 1 corresponds to plant line 5.1A7, lane number 2 corresponds to plant line 5.1A11, and lane number 3 corresponds to plant line 5.1A18. Molecular weight markers (dots on the right hand side) are 50 kDa (upper dot) and 40 kDa (lower dot). His-antibody was used as the probe.

Figure 14 shows a Western blot of HYB-oleosin expressing *Brassica campestris* lines. Lane number 1 corresponds to plant line 17.1237, lane number 2 corresponds to plant line 17.1238, lane number 3 corresponds to plant line 17.1240, lane number 4 corresponds to wild-type (untransformed) control, lane number 5 corresponds to plant line number 17.20c8, lane number 6 corresponds to plant line number 17.20c11, lane number 7 corresponds to plant line 17.20c20. Molecular weight markers (lane 8, dots on the right hand side) are 50 kDa (upper dot) and 40 kDa (lower dot). His-antibody was used as the probe.

Figure 15 shows a Western blot of Nap-oleosin expressing *Brassica campestris* lines. Lane number 2 corresponds to plant line 14.17C2, lane number 3 corresponds to plant line

14.17C15, lane number 4 corresponds to wild-type (untransformed) control, lane number 5 corresponds to plant line 14.17C16, lane number 6 corresponds to plant line 14.17C18, lane number 7 corresponds to plant line 14.12N, lane number 8 corresponds to plant line 14.12N8. Molecular weight markers (lanes 1 and 9) are 50 kDa (upper dot) and 40 kDa (lower dot). His-antibody was used as the probe.

Detailed Description of the Invention

Definitions

The terms used in the present invention have the meaning they usually have in the fields of plant biotechnology, protein chemistry and feed formulation. Some terms in the present invention are, however, used in a broader or somewhat different manner. Therefore, some of the terms are defined in more detail below.

"Recombinant nucleotide sequence construct" or "construct" means a DNA construct, an expression cassette or a transformation vector. It comprises linear or circular end-to-end linked nucleotide sequences optionally inserted in a plasmid. The constructs are intermediate constructs expressing a reporter-protein, which enables selection (identification) of effective constructs from which the reporter-sequence is then removed to provide transformation constructs for transforming any suitable plant species. In the present invention the construct comprises regulatory sequences, a nucleotide sequence encoding a carrier-protein and a nucleotide sequence encoding a reporter protein. At least one nucleotide sequence cassette, but preferably from one to six cassettes, each cassette comprising at least two, preferably five, more preferably ten, most preferably ten to fifteen of the selected amino acid codons are fused between the nucleotide sequences encoding the carrier- and the reporter-protein.

An "amino acid cassette" in the present invention means an insert comprising a continuous nucleotide sequence having at least two, preferably four, more preferably ten, most preferably ten to eighty codons or triplets encoding an optimal number of the desired amino acid. Preferably, the optimal number of codons is one that provides a stable amino protein enriched with desired amino acid. Typically the number of amino acid residues, which can be stably

attached as an extension is from ten to eighty amino acid residues. The term "amino acid cassette" is also used for the polyamino acid chain, which is stably attached to a carrier-protein. The number of amino acid residues, which can be attached to the carrier-protein, is determined by randomly inserting nucleotide sequence cassettes comprising different numbers of amino acid codons into the construct and screening in a cell free translation system to confirm correct codon translation.

"Targeted accumulation" means that the amino acid-enriched protein is expressed in certain selected plant tissues or organs or is transported to said desired tissues or organs. This can be achieved by selecting carrier-proteins, which take part in the "intracellular trafficking pathways" of the plant and thereby transport the expressed protein to certain organs and tissues and thereby enable accumulation of the expressed product for example in the cell walls or cell membranes in seed. For example, expression of amino acid-enriched carrier-proteins in seeds of transgenic *Brassica* species targeted to the cell walls of the seed cells are anchored to the oilcake remaining after pressing out the oil. The oil cake, provided with improved properties, i.e. increased amino acid content, is a useful ingredient in composition for the feed industry. In a similar manner, the amino acid-enriched protein can be expressed in mesophyll tissue of lettuce, and used directly as human nutrient source.

The selected amino acid can be any amino acid, but preferably they are the eight essential amino acids required by human beings or livestock. Selected amino acids are, for example, histidine, cysteine, methionine, glycine, tryptophan, lysine, alanine, valine, leucine, isoleucine, proline, phenylalanine, tyrosine, serine, threonine, arginine, aspartate, glutamate, asparagine and glutamine. Any amino acid-enriched proteins can be prepared as exemplified in the present invention by the preparation of histidine-enriched protein. The amino acid-enriched protein is obtained by providing a construct comprising a nucleotide sequence encoding a carrier-protein with an extension of an optimal number of one or more of the desired amino acid codons. In the preferred embodiment of the present invention the amino acid codons are fused (situated or placed) between the nucleotide sequences encoding the carrier- and the reporter-protein in such a way that the additional amino acid residues do not disturb the normal biological functions of the carrier-protein, for example, taking part in a secretory intracellular trafficking pathway.

"Carrier-protein" in the present invention means a protein, which can be stably extended by a cassette comprising any amino acid codons, without disturbing the normal biological functions of the corresponding native unmodified protein. Nucleotide sequences encoding carrier-proteins useful in the present invention are selected from plant proteins that use secretory intracellular trafficking pathways, which enable accumulation of the amino acid-enriched protein in the cell walls or membranes of desired plant organs such as seeds and leaves. The carrier-proteins of the present invention can be derived from genes encoding three major seed proteins. These genes are useful model proteins for carrying out genetic engineering with seed proteins especially from the family *Cruciferae*. Yet, genetic engineering can be applied in any other plant species if necessary. Three potentially applicable carrier-proteins are the seed proteins cruciferin (500 amino acid residues), napin (165 amino acids) and oleosin (165 amino acids), but other proteins can be used as well. All the above mentioned proteins are present in native seeds of *B. campestris* and are poor in histidine content: cruciferin contains only 9 His residues per 500 amino acids, napin contains only 2 His residues per 165 amino acids oleosin does not contain His at all.

In the present invention oleosin has been used as a model carrier-protein but other carrier proteins can be used in similar manner. For example, two related proteins, caleosin and steroleosin are suggested as potentially useful due to the similar way they accumulate in oil bodies and cell walls of seeds. The protein oleosin is reviewed e.g. in (i) Murphy 1996. TIBTECH 14. 206-213; (ii) Methods in Mol.Biol. vol.44: Agrobacterium protocol. Eds. K.M.A.Gartland and (iii) M.R.Daey, Humamana Press Inc. Totowa, NJ; and Brassica Oilseeds: Production and Utilization. Eds D.S.Kimber and D.I. McGregor. Cab International. 1995). Oleosin (Ole) was shown to be a useful carrier-protein, because it is a hydrophobic protein, has a relatively small size and is a component of the membranes surrounding the storage oil-bodies of *B. campestris* seeds.

A potential disadvantage when using the oleosin nucleotide sequence for *B. campestris* transformation is connected with the probability of inducing in transgenic plants, post-transcriptional gene silencing due to nucleotide sequence homology between the transgene and the endogenous gene. This probability was successfully lowered, partly by the fact that the inventors used the Ole gene from *Arabidopsis thaliana,* not from *B. campestris.* In order

to overcome the potential risk of gene silencing, a gene of non-plant origin (namely, 30K TMV MP) was used and assessed in a parallel series of experiments.

An advantage of using MP was that this gene has no sequence homology to endogenous genes of *B. campestris,* but encodes a protein, which could satisfy all the requirements for amino acid-enriched protein expression in the seeds mentioned above. In particular, the 30K TMV MP is a hydrophobic membrane-bound protein (Reichel and Beachy, 1998, PNAS95, 11169-11174) and it has been well documented that TMV MP is targeted to cell wall (CW) fraction and is accumulated at plasmodesmata (PD) (Oparka et al., 1997, Plant J. 12, 781-789). However, TMV MP never accumulates in seeds. To overcome the problem the carrier-protein TMV MP was linked to regulatory sequences, regulating the expression in such a way that accumulation in seeds is ensured. This is mediated by using the appropriate promoter(s) as demonstrated in the examples. For instance, it was found that the said protein, TMV MP, which is not normally expressed in seed, could be introduced into seeds when expressed from a napin promoter.

"Regulatory sequence" means nucleotide sequences, which regulate the transcription and expression of the structural nucleotide sequences either by down-regulating or up-regulating transcription and expression. Regulatory sequences comprise promoters, enhancers, signal sequences, terminators, etc. Preferred promoters are relatively short, organ- and/or tissue-specific transcriptional promoters driving the transcription of the chimeric nucleotide sequence during different stages of morphogenesis and particularly during embryogenesis. Other potentially useful promoters are plant virus-derived promoters from AMV and CaMV (19S), and those that can be regulated either by environmental effects (e.g. heat) or under specific abiotic (e.g. salicylic acid-responsive) or biotic (pathogen) stress conditions.

The preferred "transcription promoters" in the present invention are a napin promoter, a 35S promoter and/or a chimeric ("hybrid", HYB) promoter. Naturally, promoters can include any other operating regulatory sequences, particularly promoters enabling accumulation of amino acid-enriched proteins in selected, preferred tissues, e.g. in oil bodies and cell walls of seeds.

In the present invention the exemplified promoters are the napin promoter from *Arabidopsis*

thaliana, the 35S promoter of the cauliflower mosaic virus (CaMV) and/or the chimeric ("hybrid", HYB) promoter, which comprises the entire napin promoter coupled with the enhancer sequence of the CaMV 35S promoter.

The napin promoter controlling the expression of the napin (NAP) gene is regulated during embryogenesis and is switched on after flowering. The nucleotide sequence of NAP promoter from *A. thaliana* has been reported (Rask et al. 1998. J. Plant. Physiol. 152, 595-599). An important advantage of the NAP promoter in the present invention is the fact that it is rather short (152 bp) in comparison to the promoter of Ole gene which is 2.1 kb.

The 35S promoter of cauliflower mosaic virus (CaMV) has been shown to be active in cells of *B. campestris* (Harpster et al., 1988. Mol. Gen. Genet. 212, 182-190). The 35S promoter is also active in embryos and the detection of transgenic *B. campestris* plants would be facilitated by testing the product of amino acid-enriched recombinant protein (Ole or TMV MP) in mature leaves before flowering.

A "reporter sequence" or "nucleotide sequence encoding a reporter-protein" means a nucleotide sequence used to design a construct, which enables the selection of correct constructs enabling stable transformation and expression of stable amino acid-enriched proteins with intact biological functions capable of accumulating in the desired plant tissues. In this case, the nucleotide sequence encoding the reporter-protein enables easy, accurate and unambiguous identification of transformation. The reporter sequence enables demonstration of the tissues in which the promoter is activated and the conditions under which the promoter is active. In the preferred embodiments of the present invention the reporter-protein is a visible or detectable, preferably fluorescent protein, selected from a group consisting of green fluorescent protein, β-glucuronidase, obelin or luciferase.

The following selection systems are used in the present invention:

A cell-free translation system enabled the selection of the optimal number of codons that could be translated correctly.

A transient expression assay was used to ensure protein expression and that the expressed protein had intact biological function(s) as compared to the native unmodified protein, and was accumulating in the targeted organ. This was facilitated by visual observation of the reporter-protein.

Confocal laser scanning microscopy allowed early detection of transgenic plants expressing amino acid-enriched proteins in-frame with the nucleotide sequence encoding the reporter-protein.

General Description of the Invention

The primary objective of the present invention was to elaborate methods providing a means to increase the content of selected amino acids in plant cells tissues, particularly in seeds in the family *Cruciferae*. With the method of the present invention, stable amino acid-enriched protein was successfully expressed during embryogenesis of transgenic *Brassica campestris* plants and accumulated in the seeds. With the method and constructs of the present invention the said recombinant amino acid-enriched protein was tightly bound in seed cells to membranes or cell walls (CW) and, therefore, it could be retained in seed remnants, anchored to the oilcake, after the oil was pressed out. This enabled oil to be recovered with subsequent recovery and use of the remains, i.e. the oil cake, which with the improved properties, i.e. increased amino acid content, provides a useful ingredient in an amino acid-enriched composition for preparing animal feed. When the extra amino acid, for example, histidine, is in cow fodder, and is bound into normal, native proteins as in the present invention, it functions as a natural amino acid and can be utilized normally in cow metabolism. Therefore, an amino acid-enriched, for example histidine-enriched, protein is a useful ingredient in fodder for cows, functioning as a booster of milk production. Enrichment of other useful amino acids, for example, in edible lettuce leaves would allow the use of such amino acid-enriched protein(s) as a direct human food source with no further processing.

Nucleotide sequences encoding reporter-proteins, such as green fluorescent protein (GFP), positioned downstream of the selected, desired amino acid-codon enriched sequence(s) allowed easy detection of gene expression and intracellular localization of the expressed

proteins in transgenic plants.

Therefore, the present invention pertains to the methods for producing transformed plants and plant tissues that are capable of expressing a high level of amino acid-enriched stable proteins, which are in particular localized in different cell compartments including seed membraneous oil bodies and cell walls.

The method of the present invention comprises transformation of plants with constructs enabling expression of stable proteins enriched with any amino acid in targeted plant tissues. The construct for stable transformation comprises regulatory sequences including an organ- and/or tissue-specific transcriptional promoter driving the transcription of the gene(s) of interest during different stages of morphogenesis, particularly during embryogenesis operably linked to: (a) a nucleotide sequence encoding a carrier-protein without a termination codon; and (b) a nucleotide sequence comprising at least one cassette including at least two codons encoding the desired amino acid residues fused in frame with the nucleotide sequence encoding the carrier protein.

The desired constructs for stable transformation of plants expressing stable amino acid-enriched protein are obtainable by a method, wherein the above defined construct without a termination codon is fused in frame with a nucleotide sequence encoding a reporter-protein.

The reporter-protein enables easy and accurate selection in a cell-free system of constructs providing expression of stable amino acid-enriched carrier-proteins having a detectable reporter-protein combined with a transient expression system that enables verification that the biological function(s) of the carrier-protein are not compromised. In the case of oleosin, the intact biological function means that the reporter-protein is expressed in oil bodies of seeds, particularly in the membranes and cell walls of seed cells.

Those constructs exhibiting normal biological function(s) of the carrier-protein as visualized by the reporter-protein were selected and the nucleotide sequence encoding the reporter-protein was removed. Accordingly, the carrier-protein and the codons encoding the desired amino acids are fused in frame.

The satisfactory constructs were selected and transferred into *Agrobacterium,* and positive (construct-containing) clones were selected by Southern blot hybridization and used for transformation of crop plants, particularly in the family *Cruciferae.* Transformed crop plants expressed stable amino acid-enriched proteins and accumulated the said proteins in the targeted plant tissue(s).

The regulatory sequences, including the transcriptional promoter, were selected among examples including tissue- and/or organ-specific transcription promoters, which enable targeting of mRNA synthesis to leaf, seed, or any other plant organs and, thereby, provide targeted accumulation of amino acid-enriched protein(s). In order to enable the selection of satisfactory constructs, the amino acid codon-enriched sequence coding for selected, desired amino acid residues is placed in frame between the amino acid carrier-protein gene and the nucleotide sequence encoding (preferably) a fluorescent reporter-protein, e.g. green fluorescent protein (GFP), ß-glucuronidase and/or luciferase.

The preferred crop plants of the present invention belong to the family *Brassicaceae,* wherein amino acid-enriched, particularly histidine-enriched proteins, based on their ability to be anchored to and localized in the membranes of oil-bodies after expression, provide targeted accumulation. Based on the same principles, other constructs enabling accumulation of any other amino acid-enriched proteins in any other desired plant species and their tissues can be obtained.

The above objectives are accomplished by transforming plants with constructs comprising nucleotide sequences encoding carrier-proteins selected from relatively small plant proteins functioning in selected secretory intracellular trafficking pathways enabling accumulation in desired, targeted plant organs, e.g. leaf membranes, seed membranes, cell walls of leaves or seed cells. This is exemplified in the present invention by two distinct carrier-proteins, the oleosin protein and the TMV protein comprising a polyamino acid tail fused in frame to said natural carrier-protein subsequently accumulating in desired targeted tissues of said transgenic plants. Particularly preferred nucleotide sequences coding for the carrier-proteins are those of oleosin, which is capable of accumulating in leaves or seed membranes and TMV MP, which can accumulate in cell walls of leaves or seed cells, when transcription is driven

by the appropriate promoter.

The nucleotide sequence encoding one or more amino acid in a cassette is enriched with the selected amino acid codons and the nucleotide sequence cassette is preferably centrally located such that it does not disturb the normal biological function(s) of the carrier-protein. If the amino acid codons are placed in the N-terminal end or somewhere in the middle of carrier-protein, the amino acid-enriched protein may not accumulate in the targeted organ as desired. The nucleotide sequence encoding the reporter-protein is fused in the same translational frame of the extended nucleotide sequence enriched with the selected amino acid codons, lacking the termination codon. The optimal number of amino acid codons is about ten to eighty and is determined by checking correct protein translation in a cell-free translation system.

Therefore, in a preferred exemplified embodiment of the invention the nucleotide sequence encoding the carrier-protein is positioned in the 5'-proximal end of the nucleotide sequence and the reporter nucleotide sequence is positioned in the 3'-terminal end. The carrier-proteins or nucleotide sequence of interest in the present invention are OLE or 30K TMV MP.

The OLE used in the examples of the present invention is a nucleotide sequence corresponding to the chromosomal gene of *Arabidopsis thaliana* coding for the seed protein oleosin (Ole), which is a component of membranes of oil-bodies in members of *Cruciferae* family. 30K MP is a nucleotide sequence derived from the genomic RNA of tobacco mosaic virus U1 (TMV U1). The gene encodes a non-structural hydrophobic 30K protein responsible for the movement of the viral genome from cell to cell in an infected plant (movement protein, MP) through plasmodesmata (PD). The 30K MP is targeted to and accumulates in cell walls and PD, and is expressed in plants transgenic for the MP gene.

The optimal number of amino acid codons is determined by randomly transforming constructs with nucleotide sequence encoding the carrier-protein with extensions of different sizes, comprising different numbers of amino acid codons, into plants and selecting transgenic plants expressing proteins with undisturbed biological functions as compared to the native unmodified protein, assessed using a transient expression assay.

Functional constructs and successful transformation can easily be demonstrated by the expression and intracellular localization of the expressed proteins, particularly the reporter-protein, in transgenic plants.

In the present invention, genetic engineering methods are accordingly improved and the new principle of selection allows all stages of transgenic plant production to be controlled and permits the rapid identification and selection of any amino acid-enriched chimeric gene of interest that has optimal codon content.

In the present invention the method for producing amino acid-enriched proteins in transgenic plants involves the preparation of one or more of the constructs described above.

The preferred selection method comprises expression of a fluorescent recombinant amino acid-enriched protein from constructs such as (Ole-polyamino acid-GFP or TMV MP-poly-amino acid-GFP) in a cell-free translation system. The cell-free translation system used in this invention is a method that unequivocally identifies correct constructs and expression of the recombinant amino acid-enriched proteins. Additionally, detection of fluorescence, provided by GFP, or other reporter-proteins, can be carried out using confocal laser scanning microscopy of seeds and/or leaves. The use of confocal laser scanning microscopy provides a new convenient tool for early detection and selection of transgenic plants expressing amino acid-enriched proteins in-frame with the reporter-protein, e.g. GFP.

The reason for using a reporter nucleotide sequence insert was that this protein could be used as a marker to confirm that the constructs and the expression of the amino acid-enriched carrier-protein, e.g. oleosin-polyamino acid-GFP or TMV MP-poly-amino acid-GFP, from recombinant sequences were correct. Fluorescence, particularly GFP-fluorescence can be monitored by confocal laser scanning microscopy that allows detection of Ole-polyamino acid and MP-polyamino acid expression *in vivo* as well as determination of their expression levels.

Plant expression vectors comprise nucleotide sequences encoding carrier-proteins, such as Ole and TMV MP fused to polyamino acid-coding sequences of varying length and carrying

GFP gene positioned in the 3'-terminal end. To construct the plant expression vectors, said constructs were fused with sequences of varying lengths enriched with desired amino acid codons, e.g. for histidine the codons CAC and CAU, for cysteine and methionine the codons TGT, TGC and ATG, for glycine the codons GGA, GGT, GGC, and GGG, and for lysine the codons are AAA and AAG. These constructs lacked a termination codon. The resulting constructs contained randomly selected number but at least two amino acid codons, inserted as at least one cassette, or for example by adding one, two, four, six or eight cassettes, each comprising 14 histidine codons resulting in 14 (x), 28(2x), 56(4x) and 112 (8x) His-codons downstream of the gene (Ole or MP). It was found that 8x (112 aa) were unstable producing deletion variants (mini-plasmids); some of the 6x (84 aa) clones were unstable. Thus, for further work clones 2x, 4x and 6x were fused with the 3'-proximal GFP gene with 3' termination sequences. In precisely the same way it is possible to determine the optimal number of any amino acid codons, which can be inserted to any selected carrier-protein. It is, however, self-evident that cassettes comprising any number of the desired amino acid codons can be used.

Constructs, such as those shown in Figure 1, comprising regulatory sequences, particularly the promoters NAP, 35S and HYB were operably linked to a carrier-protein, particularly Ole and/or TMV MP. All of the constructs comprised a nucleotide sequence encoding a reporter-protein, particularly preferred is the green fluorescent protein (GFP), as a marker for selection. The GFP gene was fused in frame to the amino acid-tail, which was of varying length including one or more cassettes comprising at least two amino acid codons encoding the desired amino acids. Particularly, multiples of the cassettes were used (2x, 4x, 6x). Fluorescence of the reporter-protein in transient expression assays was used in preliminary experiments to identify and select the best functional constructs. The transient expression assays included micro-projectile bombardment of crop plant embryos as exemplified by experiments involving particle bombardment of B. campestris embryos and Nicotiana benthamiana leaves, and shown in Table 1. The assays demonstrated that the NAP promoter was active in embryos, whereas the HYB promoter was efficient in both embryos and epidermal leaf cells. These data formed the basis for selection of four (of the said available eighteen constructs, Figure 1.) for subsequent work with transformations (Figure 2).

Each gene construct described above was transferred into *Agrobacterium tumefaciens*; transformed clones were selected by Southern blot hybridization. Each of the transformed strains of *Agrobacterium* was then used to transform oilseed spring rape *(B. campestris)*. Subsequent selection of transgenic plants expressing the amino acid-enriched carrier-proteins (Ole-polyamino acid or MP-polyamino acid) was performed by:

(a) PCR, with primers specific for the selected amino acid codon-containing Ole and TMV MP genes
(b) Western blotting with antibodies raised against oleosin and 30K TMV MP
(c) Detection of GFP fluorescence in leaves, flowers, embryos.

Using the above experimental approach, when clones expressing proteins with undisturbed function have been found, the reporter nucleotide sequences can be removed from the corresponding construct(s) and plants transformed with the said functionally proven, construct(s).

In embodiments of our invention, a stable amino acid-enrichment in seed proteins could be demonstrated.

The invention is described in more detail in the following examples. The production of amino acid-enriched protein is exemplified by histidine, methionine, cysteine, glycine and lysine but it is self-evident that the invention can be applied to any other desired amino acid(s) by inserting cassettes comprising the selected desired amino acid codons for any specified purposes.

EXAMPLE 1

Chromosomal DNA isolation and PCR

Chromosomal DNA was isolated as follows. 300 μg of *A. thaliana* leaf tissue was homogenized in a mortar in liquid nitrogen to a fine powder. Then 3 ml of 100 mM Tris-HCl; pH 8.0, 500 mM NaCl, 50 mM EDTA was added, with further grinding after the addition of

600 μl of 10% SDS and 500 μl of 20% polyvinylpyrrolidone (average mwt 360,000). The mixture was transferred to a polypropylene tube and incubated at 65°C for 10 min. Then 450 μl of ice-cold potassium acetate was added to the solution and the mixture was gently shaken by tube inversion, incubated on ice for 30 min and centrifuged for 10 min at 8000 rpm, 4°C. The supernatant was extracted with phenol/chloroform (1:1), and the DNA was precipitated using iso-propanol. The precipitated DNA was washed with 70% ethanol, and air-dried prior to being dissolved in 10 mM Tris-HCl; pH 8.0. The PCR reaction mixture was composed of 1 x PCR buffer (10 x buffer: 500 mM KCl, 100 mM Tris-HCl; pH 9.0 at 25°C and Triton X-100) with 1.5 mM $MgCl_2$, 1μl of Taq polymerase (5U/μl), 0.2 mM of each dNTPs, primers at 0.4 μM and 3 μg of A. thaliana genomic DNA in 25μl reaction volume. The template was denatured with heating for 3 min at 95°C, and 30 cycles of PCR were carried out with iCycler™ (Bio-Rad) thermal cycler with denaturation at 95°C for 1 min, primer annealing at 65°C for 1.5 min, primer extension at 72°C for 2 min, with a final elongation step after 30 cycles at 72°C for 10 min. The PCR product of expected size (776 bp) was isolated from a 1% agarose gel after electrophoresis of the final PCR reaction mix, purified using a gel slice kit (Qiagen), and then treated with 1 unit of T4 DNA polymerase in 1X T4 DNA polymerase reaction buffer (50 mM NaCl, 10 mM Tris-HCl; pH 7.9, 10 mM $MgCl_2$, 1 mM dithiothreitol) supplemented with 100 mM dNTP at 14°C for 15 minutes prior to cloning.

EXAMPLE 2

Synthesis and cloning of a histidine codon-enriched DNA sequence

To construct a histidine codon-enriched DNA fragment, three chemically synthesized ssDNA segments were used:

H-P-1
(5'-GCGCCTCGAGTTCACCATCACCATCACCATCACGGGCACCATCAC) (SEQ ID NO:1),

H-P-2
(5'-CATCACCATCACCATGG) (SEQ ID NO:2)

and

H-M

(5'-CCGGATCCTAAAGTCGACCATGGTGATGGTGATGGTGATGGTGCC) (SEQ ID NO:3).

To obtain the dsDNA fragment carrying the His codon-enriched sequence, oligonucleotides H-M and H-P-2 were allowed to anneal in 1X AMV reverse transcriptase (RT) reaction buffer (50 mM Tris-HCl; pH 8.3, 50 mM KCl, 10 mM $MgCl_2$, 0.5 mM spermidine, 10 mM DTT) for 30 min at room temperature, and chain elongation reaction was carried out in the presence of 1X AMV RT buffer, 1 mM dNTPs and 1 unit of AMV RT for 45 min at 37°C. The reaction product was separated from non-annealed oligonucleotides by agarose gel electrophoresis, purified, and annealed with the oligonucleotide H-P-1 in 1X DNA polymerase I large (Klenow) fragment buffer (50 mM Tris-HCl; pH 7.2, 10 mM $MgSO_4$, 0.1 mM DTT) for 30 min at room temperature. The chain elongation reaction was performed with 5 units of Klenow polymerase in the presence of 25 mM of each dNTPs in 1X Klenow buffer for 30 min at 37°C. The dsDNA fragment of expected size (79 bp) was excised from a agarose gel (after appropriate electrophoretic separation), purified digested with XhoI and BamHI, and cloned into the similarly digested cloning vector pGEM-7Zf(+). After restriction analysis and sequencing, clone pGEM-His-24 was selected for further manipulations. This clone contained a DNA segment that could potentially code for a 19 amino acid-long peptide containing 14 His residues (Fig. 3a.). The sequence of this DNA fragment was flanked by *Xho*I and *Bam*HI sites and contained also a *Sal*I restriction site designed for subsequent cloning steps (Fig. 3a.).

EXAMPLE 3

Synthesis and cloning of a cysteine and methionine codon-enriched DNA sequence

C-M-P-1

(5' CACCTCGAGTATGTTGTTGCATGTGCATGTGCTGTTGCATGTCGACAAAC 3') (SEQ ID NO:4)

C-M-P-2

(5' GTTTGTCGACATGCAACAGCACATGCACATGCAACAACATACTCGAGGTG 3') (SEQ ID NO:5)

To obtain the dsDNA fragment carrying the Cys/Met codon-enriched sequence, oligonucleotides C-M-P-1 and C-M-P-2 were allowed to anneal in 1X DNA polymerase I large (Klenow) fragment buffer (50 mM Tris-HCl; pH 7.2, 10 mM $MgSO_4$, 0.1 mM DTT) for 30 min at room temperature. The dsDNA fragment of expected size (50 bp) was excised from a agarose gel (after appropriate electrophoretic separation), purified, digested with *Xho*I and *Bam*HI, and cloned into the similarly digested cloning vector pGEM-7Zf(+). After restriction analysis and sequencing, clone pGEM-Cys/Met-10 was selected for further manipulations. This clone contained a DNA segment that could potentially code for 16 amino acid-long peptide containing 10 Cys/Met-amino acid residues. The sequence of this DNA fragment was flanked by *Xho*I and *Bam*HI sites and contained also a *Sal*I restriction site designed for subsequent cloning step (Fig. 3b).

EXAMPLE 4

Synthesis and cloning of a glycine codon-enriched DNA sequence

GL-P-1

(5'-GCGCCTCGAGTTGGTGGAGGTGGAGGCGGTGGAGGTGGCGTCGACAAATGG ATCCCC-3') (SEQ ID NO:6)

GL-P-2

(5'-GGGGATCCATTTGTAGACGCCACCTCCTCCACCGCCTCCACCTCCACCAACTC GAGGCGC-3') (SEQ ID NO:7)

To obtain the dsDNA fragment carrying the glycine codon-enriched sequence, oligonucleotides GL-P-1 and GL-P-2 were allowed to anneal in 1X DNA polymerase I large (Klenow) fragment buffer (50 mM Tris-HCl; pH 7.2, 10 mM $MgSO_4$, 0.1 mM DTT) for 30 min at room temperature. The dsDNA fragment of expected size (60 bp) was excised from agarose gel (after appropriate electrophoretic separation), purified, digested with *Xho*I and *Bam*HI, and cloned into the similarly digested cloning vector pGEM-7Zf(+). After restriction analysis and sequencing, clone pGEM-Gly-9 was selected for further manipulations. This clone contained a DNA segment that could potentially code for a 19 amino acid-long peptide containing 9 Gly-amino acid residues. The sequence of this DNA fragment was flanked by *Xho*I and *Bam*HI sites and contained also a *Sal*I restriction site designed for subsequent cloning steps (Fig. 3c.).

EXAMPLE 5

Synthesis and cloning of a lysine codon-enriched DNA sequence

L-P-1
(5'-GCGCCTCGAGTTAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGGT CGACAAATGGATCCCC-3') (SEQ ID NO:8)

L-P-2
(5'-GGGGATCCATTTGTCGACCTTTTTCTTTTTCTTTTTCTTTTTCTTTTTCTTTTTAA CTCGAGGCGC-3') (SEQ ID NO:9)

To obtain the dsDNA fragment carrying the lysine codon-enriched sequence, oligonucleotides L-P-1 and L-P-2 were allowed to anneal in 1X DNA polymerase I large (Klenow) fragment buffer (50 mM Tris-HCl; pH 7.2, 10 mM $MgSO_4$, 0.1 mM DTT) for 30 min at room temperature. The dsDNA fragment of expected size (66 bp) was excised from agarose gel (after appropriate electrophoretic separation), purified, digested with *Xho*I and *Bam*HI, and cloned into the similarly digested cloning vector pGEM-7Zf(+). After restriction analysis and sequencing, clone pGEM-Lys-12 was selected for further manipulations. This clone contained a DNA segment that could potentially code for a 22 amino acid-long peptide

containing 12 lysine amino acid residues (Fig. 3d.) The sequence of this DNA fragment was flanked by *Xho*I and *Bam*HI sites and contained also a *Sal*I restriction site designed for subsequent cloning (Fig. 3d.).

EXAMPLE 6

Isolation of the oleosin gene from the *Arabidopsis thaliana* chromosomal DNA using PCR, and subsequent cloning

To clone the oleosin gene from *Arabidopsis thaliana* chromosomal DNA, two oligonucleotides were chemically synthesized:

OLE-P

(5'-AAAACCATGGCGGATACAGCTAGAGGAACCCATC) (SEQ ID NO:10) and

OLE-M

(5'-GGGGCCATGGGAGTAGTGTGCTGGCCACCACGAGTAC) (SEQ ID NO:11).

OLE-P and OLE-M were used as the primers for PCR of genomic DNA from *Arabidopsis thaliana* (example1).

The resulting PCR was fragment was blunt-ended, and cloned into *Sma*I site of pGEM-3Zf(+). The recombinant clones were verified by restriction analysis and sequencing. For subsequent work, clone pOLE4H was selected in which the oleosin gene differed in two nucleotide positions from the published sequence with no changes in the encoded amino acid sequence. To facilitate construction of oleosin fusions, clone pOLE4H was modified to replace the terminator codon with an appropriate restriction site (*Xho*I). To achieve this, the oligonucleotide :

OLE-3' XhoI

(5'- GCGCCTCGAGAAGTAGTGTGCTGGCCACCAC) (SEQ ID NO:12)

was chemically synthesized, and used as primer for PCR of pOLE4H, and subsequently cloned into *Xho*I site of pGEM7Zf(+). After the cloning of the pOLE4H sequence using this primer, the entire sequence of the resulting clones was verified by sequencing to ensure that all of the oleosin gene coding properties were retained. Based on this analysis, clone pOLE4/11 was selected. pOLE4/11 contained a wild-type oleosin gene without a terminator codon, flanked by *Nco*I and *Xho*I restriction sites suitable for subsequent cloning steps.

EXAMPLE 7

Isolation of the 30K movement protein gene from tobacco mosaic virus (30K TMV MP) genomic RNA using RT-PCR and subsequent cloning

To clone 30K gene from TMV genomic RNA, two specific oligonucleotide primers

N30K
(5'-GCGGAATTCCCATGGCTCTAGTTGTTAAAGG) (SEQ ID NO:13) and

C30K
(5'-AGACCTCGAGGAAACGAATCCGATTCGGCGAC) (SEQ ID NO:14)

were chemically synthesized; these primers contained an EcoRI restriction site and XhoI restriction site, respectively. The first strand cDNA was synthesized from TMV genomic RNA using 20 nM of C30K primer, 1 mM dNTPs and 1 unit of AMV in 1X AMV reverse transcriptase (RT) reaction buffer (50 mM Tris-HCl; 8.3, 50 mM KCl, 10 mM $MgCl_2$, 0.5 mM spermidine, 10 mM DTT) for 45 min at 37°C. After that 5 μl of reaction mix was subjected to PCR reaction in 1 x PCR buffer (10 x buffer: 500 mM KCl, 100 mM Tris-HCl; pH 9.0 at 25°C and Triton X-100) with 1.5 mM $MgCl_2$, 2μl of *Taq* polymerase (5U/μl), 0.2 mM of each dNTPs, primers at 0.4 μM in 25μl reaction volume. The template was denatured with heating for 3 min at 95°C, and 28 cycles of PCR were carried out with iCycler (Bio-Rad) thermal cycler with denaturation at 95°C for 1 min, primer annealing at 68°C for 1.5 min,

primer extension at 72°C for 2 min, with a final elongation step after 30 cycles at 72°C for 10 min. The resulting DNA product was cleaved by *Eco*RI and *Xho*I and cloned into *Eco*RI-*Xho*I-digested pGEM-7Zf(+). Following restriction analysis of the recombinant clones and sequencing, plasmid pGEM-30K was selected for further work. In this plasmid, the termination codon of the TMV 30K gene was replaced by a *Xho*I site for subsequent fusion with the His, Cys/Met, Gly and Lys codon-enriched sequence (examples 2-5).

EXAMPLE 8

Particle bombardment

Particle bombardment was performed using the rupture disk method with a high-pressure helium-based apparatus PDS-1000 (Bio-Rad) as described in Morozov et al. (1997). Tungsten particles were prepared by vortexing of 60 mg in 70 % ethanol for 3-5 minutes, followed by incubation on a bench for 15 min. Mixture was pelleted by short spinning and the supernatant was removed. Sterile water was added onto pellet and vortexed for 1 min. Particles were allowed to settle for 1 minute and spinned for 2 s. The supernatant was removed. This was repeated three times. Sterile 50 % glycerol was added to bring particle concentration to 60 mg/ml. For each particle bombardment, DNA was precipitated onto tungsten particles (M-20, 1.3μ) with calcium chloride and ethanol. 5-10 μg of plasmid DNA, 50 μl of $CaCl_2$ (2.5 M) and 20 μl of spermidine (0.1 M) were mixed and vortexed for 2-3 min, allowed to settle and spinned for 2 s. Supernatant was removed and 140 μl of 70 % ethanol was added onto surface of pellet, removed and 100 % ethanol added and removed without disturbing the pellet, and repeated. 6 μl of suspension was pipetted onto macrocarrier and used for bombardment. A detached leaf of *Nicotiana benthamiana* (15-30 mm size) was placed in the center of a plastic Petri dish and bombarded on a solid support at a target distance of 7 cm. Bombardment was done with a pulse of 1350 kPa helium gas in a vacuum chamber. Inoculated leaves were analyzed 24 hours after particle bombardment. GFP fluorescence was monitored using a confocal laser scanning imaging system MRC-1024 (Bio-Rad).

EXAMPLE 9

Construction of an *E. coli* strain over-expressing the C-terminal hydrophilic terminus of oleosin for subsequent rabbit immunization and obtaining oleosin-specific antibodies

Most of the oleosin N-terminal sequence region is hydrophobic, whereas the amino acid sequence of the oleosin C-terminal portion encoded by the second exon of the oleosin gene represents hydrophilic sequence that is suitable for expression in *E. coli* and using as the antigen in immunizations. Therefore, the C-terminal oleosin region was selected for this purpose. The second exon of the oleosin gene was amplified in a PCR reaction using the plasmid pOLE4/11 as the template and the oligonucleotide primers:

OLE-EX-B
(5'-TGTGGGATCCTACGCAACGGGAGAGCACCCA) (SEQ ID NO:15) and

OLE-3'XhoI
(5'-GCGCCTCGAGAAGTAGTGTGCTGGCCACCAC) (SEQ ID NO:12)

containing *Bam*HI restriction site and *Xho*I restriction site, respectively. The resulting PCR product was digested with *Bam*HI and *Xho*I and cloned into *Xho*I- and *Bam*HI-digested pGEM-7Zf(+). After restriction analysis of the resulting clones and sequencing, clone pGEM-OLE2EX was selected for the subsequent cloning steps. To obtain an expression construct, the region of the oleosin second exon was removed from pGEM-OLE2EX by digestion with *Bam*HI and *Xho*I and cloned into the (*Bam*HI and *Sal*I digested) expression vector pQE30 (QIAGEN). In the resulting construct pQE-OLE2EX, the 3'-terminal portion of the oleosin gene was translationally fused to a leader sequence containing a 6xHis tag, which makes it possible to affinity purify the expressed product. The plasmid pQE-OLE2EX was transformed into *E. coli* strain M15 [pREP4] (QIAGEN). Induction of recombinant protein expression was carried out in Luria broth by adding IPTG to a final concentration 1 mM followed by incubation on a rotary shaker at 220 rpm for 2-4 h at 37°C. The *E. coli* cells were collected from induced and non-induced control cultures, and expression of the recombinant protein was analyzed by SDS-PAGE. For SDS-PAGE, 12.5 % acrylamide gels were

prepared. Gels were run at 100V/500mA for 45 1 h, and subsequently stained in CBB (0.05 % Coomassie Brilliant Blue, 50 % methanol, 7 % glacial acetic acid) and destained in 7 % methanol/5 % acetic acid. In the induced cultures, a major band was evident that was absent in the non-induced cultures. Mobility of this band corresponded to the expected mobility of the His-tagged C-terminal portion of the oleosin gene (8.3 kDa).

EXAMPLE 10

Fusion of histidine codon-enriched DNA sequences to the TMV 30K gene and testing the correct translation of His codon-enriched genes using an *in vitro* translation system

Fusions between His codon-enriched sequences and the 30K TMV MP gene were obtained by appropriate cloning of a combination of pGEM-His-24 containing the sequence coding for 14 His residues (= one cassette), and plasmid pGEM-30K containing the 30K TMV MP gene lacking its natural terminator codon. To obtain His codon-enriched sequences of different lengths fused to the MP gene, a stepwise cloning procedure was carried out. In the first step, the 30K MP gene was excised from pGEM-30K with *Eco*RI and *Xho*I, and the His-codon-enriched sequence from pGEM-His-24 was excised with *Xho*I and *Bam*HI. Both DNA fragments were ligated into pGEM-3Zf(+) digested with *Eco*RI and *Bam*HI (Fig. 4). The resulting clone, designated pMP-1x, had one histidine-coding sequence unit (cassette) from pGEM-His-24 fused to the TMV 30K MP gene.

Additional stepwise cloning steps performed to increase the length of the His-coding "tail" in the MP fusion were based on the pMP-1x plasmid. These clonings used the restriction sites *Xho*I and *Sal*I which have been designed in the His codon-enriched sequence of pGEM-His-24 to be situated to the reading frame of the encoded polypeptide. Together with the fact that digestions with *Xho*I and *Sal*I produce identical sticky ends, this preserved the reading frame of the resulting addition-enlarged His-containing sequence in the further cloning steps described below. To obtain a plasmid with two histidine-coding sequence units from pGEM-His-24, pMP-2x, the *Eco*RI-*Sal*I-fragment from pMP-1x was cloned into pMP-1x digested with *Eco*RI and *Xho*I, resulting in duplication of the histidine-coding sequence unit. Therefore, the resulting pMP-2x plasmid contained 24 histidine residues in the MP C-

terminal "tail". Subsequently, to obtain pMP-4x, the *Eco*RI-*Sal*I-fragment from pMP-2x was cloned into pMP-2x digested with *Eco*RI and *Xho*I; to obtain pMP-6x, the *Eco*RI-*Sal*I-fragment from pMP-4x was cloned into pMP-2x digested with *Eco*RI and *Xho*I; and to obtain pMP-8x, the *Eco*RI-*Sal*I-fragment from pMP-4x was cloned into pMP-4x digested with *Eco*RI and *Xho*I. Specifically, similar clonings were repeated to obtain plasmids carrying 56, 84 and 112 histidine residues in the MP C-terminal "tail".

Clones with 112 His codons in the C-terminal "tail" (pMP-8x series) were unstable and gave rise to deletion variants (mini-plasmids) when grown in liquid medium. Lowered stability upon repeated growth in liquid medium was also observed for some of the clones with 84 histidines (pMP-6x series). Thus, for further experiments only clones with 2, 4 and 6 repeats of the His codon-enriched sequences were selected, pMP-2x, pMP-4x, and pMP-6x.

In plasmids pMP-1x, pMP-2x, pMP-4x, and pMP-6x, the 30K MP gene fused to His-coding sequences of different lengths was placed under the control of the T7 promoter. These plasmids were linearized with *Bam*HI, prior to *in vitro* transcription with T7 RNA polymerase. Equal amounts of the resulting transcripts were translated in a wheat germ cell-free system (Amersham). It was found that transcripts of the clones with one, two and four repeats of the His-enriched sequence (trMP-1x, trMP-2x, and trMP-4x) gave rise to the products of predicted, gradually increasing molecular masses, whereas the transcript potentially coding for a protein with 6 repeats of the His-enriched sequence (trMP-6x) was translated inefficiently and gave rise to smeared products of lower molecular mass. This experiment showed that translation of the His codon-enriched sequences of more than 56 residues by plant ribosomes occurs with errors, e.g. His codons could represent “hungry” codons where ribosomes stopped and jammed. On the basis of this experiments only clones with 2 and 4 repeats were selected for further work *in vivo*.

EXAMPLE 11

Fusion of cysteine/methionine codon-enriched DNA sequences to the TMV 30K gene and testing the correct translation of Cys/Met codon-enriched genes using an *in vitro* translation system

Fusions between Cys/Met codon-enriched sequences and the 30K TMV MP gene were obtained by appropriate cloning of a combination of pGEM-Cys/Met-10 containing the sequence coding for 10 Cys/Met residues (= one cassette), and plasmid pGEM-30K containing the 30K TMV MP gene lacking its natural terminator codon, as in example 10. To obtain Cys/Met codon-enriched sequences of different lengths fused to the MP gene, a stepwise cloning procedure was carried out. In the first step, the 30K MP gene was excised from pGEM-30K with *Eco*RI and *Xho*I, and the Cys/Met-codon-enriched sequence from pGEM-Cys/Met-10 was excised with *Xho*I and *Bam*HI. Both DNA fragments were ligated into pGEM-3Zf(+) digested with *Eco*RI and *Bam*HI (Fig. 4). The resulting clone, designated pMP-Cys/Met-1x, had one cysteine/methionine-coding sequence unit (cassette) from pGEM-Cys/Met-10 fused to the TMV 30K MP gene.

To obtain a plasmid with two cysteine/methionine-coding sequence units from pGEM-Cys/Met-10, pMP-Cys/Met-2x, the *Eco*RI-*Sal*I-fragment from pMP-Cys/Met-1x was cloned into pMP-Cys/Met-1x digested with *Eco*RI and *Xho*I, resulting in duplication of the cysteine/methionine-coding sequence unit. Therefore, the resulting pMP-Cys/Met-2x plasmid contained 20 cysteine/methionine residues in the MP C-terminal "tail". Subsequently, to obtain pMP-Cys/Met-4x, the *Eco*RI-*Sal*I-fragment from pMP-Cys/Met-2x was cloned into pMP-Cys/Met-2x digested with *Eco*RI and *Xho*I; to obtain pMP-Cys/Met-6x, the *Eco*RI-*Sal*I-fragment from pMP-Cys/Met-4x was cloned into pMP-Cys/Met-2x digested with *Eco*RI and *Xho*I; and to obtain pMP-Cys/Met-8x, the *Eco*RI-*Sal*I-fragment from pMP-Cys/Met-4x was cloned into pMP-Cys/Met-4x digested with *Eco*RI and *Xho*I.

In plasmids pMP-Cys/Met-1x, pMP-Cys/Met-2x, pMP-CYs/Met-4x, and pMP-Cys/Met-6x, the 30K MP gene fused to Cys/Met-coding sequences of different lengths was placed under the control of the T7 promoter. Plasmids were linearized with *Bam*HI, and subjected to *in vitro* transcription with T7 RNA polymerase. Equal amounts of the resulting transcripts were translated in a wheat germ cell-free system (Amersham Pharmacia). It was found that transcripts of clones with one, two and four cassettes of the Cys/Met-enriched sequence (MP-Cys/Met-2x, and MP-Cys/Met-4x) gave rise to the products of predicted, gradually increasing molecular masses, As exemplified in the example 10, the increase in the length of polyamino acid sequence led to a decreased stability of transcripts. For further experiments only clones

with 2, 4 and 6 cassettes (MP-Cys/Met-2x, MP-Cys/Met-4x and MP-Cys/Met-6x) of Cys/Met codon-enriched sequences were selected.

EXAMPLE 12

Fusion of glycine codon-enriched DNA sequences to the TMV 30K gene and testing the correct translation of Gly codon-enriched genes using an *in vitro* translation system

Fusions between glycine codon-enriched sequences and the 30K TMV MP gene were obtained by appropriate cloning of a combination of pGEM-Gly-9 containing the sequence coding for 9 Gly residues (= one cassette), and plasmid pGEM-30K containing the 30K TMV MP gene lacking its natural terminator codon, as in example 10. To obtain glycine codon-enriched sequences of different lengths fused to the MP gene, a stepwise cloning procedure was carried out as in examples 10 and 11.

The resulting first clone, designated pMP-Gly-1x, had one glycine-coding sequence unit (cassette) from pGEM-Gly-9 fused to the TMV 30K MP gene. Subsequently, this cassette was "multiplied" as in examples 10 and 11; to generate clones pMP-Gly-2x, pMP-Gly-4x, pMP-Gly-6x, and pMP-Gly-8x, the *Eco*RI-*Sal*I-fragment from pMP-Gly-4x was cloned into pMP-Gly-4x digested with *Eco*RI and *Xho*I. Plasmids were linearized with *Bam*HI, and subjected to *in vitro* transcription with T7 RNA polymerase and translated in a wheat germ cell-free system (Amersham Pharmacia). As exemplified in example 10, the increase in the length of polyamino acid sequence led to a decreased stability of transcripts. For further experiments only the clones with 2 and 4 cassettes (MP-Gly-2x, and MP-Gly-4x) of glycine codon-enriched sequences were selected.

EXAMPLE 13

Fusion of Lysine codon-enriched DNA sequences to the TMV 30K gene and testing the correct translation of Lys codon-enriched genes using an *in vitro* translation system

Fusions between lysine codon-enriched sequences and the 30K TMV MP gene were obtained

by appropriate cloning of a combination of pGEM-Lys-12 containing the sequence coding for 12 Lys residues (= one cassette), and plasmid pGEM-30K containing the 30K TMV MP gene lacking its natural terminator codon, as in examples 10-12. To obtain lysine codon-enriched sequences of different lengths fused to the MP gene, a stepwise cloning procedure was carried out as in examples 10 and 12.

The resulting first clone, designated pMP-Lys-1x, had one lysine-coding sequence unit (cassette) from pGEM-Lys-12 fused to the TMV 30K MP gene. Subsequently, this cassette was "multiplied" as in examples 10, 11 and 12; to generate clones pMP-Lys-2x, pMP-Lys-4x, pMP-Lys-6x, and pMP-Lys-8x, the *Eco*RI-*Sal*I-fragment from pMP-Lys-4x was cloned into pMP-Lys-4x digested with *Eco*RI and *Xho*I. Plasmids were linearized with *Bam*HI, and subjected to *in vitro* transcription with T7 RNA polymerase and translated in a wheat germ cell-free system (Amersham Pharmacia). As exemplified in the Examples 10-12, the increase in the length of polyamino acid sequence led to a decreased stability of transcripts. For further experiments only the clones with 2 and 4 cassettes (MP-Lys-2x, and MP-Lys-4x) of lysine codon-enriched sequences were selected.

EXAMPLE 14

Isolation of the napin promoter from the *Arabidopsis thaliana* chromosomal DNA using PCR, and subsequent cloning

To clone the napin promoter from the *Arabidopsis thaliana* chromosomal DNA, two oligonucleotide primers were chemically synthesized:

NAP-P
(5'-TCTTACTCGAGTGAAACCAAATTAAC) (SEQ ID NO:16) and

NAP-M
(5'-CTTGTTAGCCATGGTTTGCTATTTGTG) (SEQ ID NO:17).

To facilitate subsequent cloning, the primer sequences contained *Xho*I and *Nco*I restriction

sites. *A. thaliana* chromosomal DNA was isolated as described for isolation and cloning of the oleosin gene (example 3). The PCR reaction was carried out as described above for the oleosin gene (example 3) except that the concentration of $MgCl_2$ in the reaction mixture was 1.5 mM. The PCR product of expected size (369 bp) was isolated from a 1% agarose gel following electrophoresis. The PCR product was purified and treated with T4 DNA polymerase in the presence of dNTP at 14°C as in example 1, and cloned into *Sma*I-digested pGEM-3Zf(+). After sequencing, clones pGEM-NAP4 and pGEM-NAP9 were selected for subsequent work.

EXAMPLE 15

Construction of expression plasmids containing the napin promoter or chimeric promoter, which comprises the entire napin promoter and enhancer region of the Cauliflower mosaic virus 35S promoter

Plant expression vectors containing the napin promoter were constructed using plasmid pRT100, which contains the 35S promoter of Cauliflower mosaic virus (CaMV 35S promoter). For construction of pNAP, in which the whole sequence of the CaMV 35S promoter was removed and replaced by the sequence of the napin promoter, the napin promoter region was excised from the plasmid pGEM-NAP4 with Ecl136II and *Nco*I and ligated into pRT100 digested with *Hind*II and *Nco*I (Fig.5). To construct pHYB in which the sequence of the napin promoter replaced the promoter region of the 35S promoter retaining the 35S promoter enhancer region upstream of the inserted napin promoter sequence, pGEM-NAP9 was digested with *Xho*I, made blunt-ended with Klenow enzyme, and digested with *Nco*I. The resulting fragment was ligated into pRT100 digested with *Eco*RV and *Nco*I.

EXAMPLE 16

Cloning of His-enriched oleosin and 30K TMV MP genes into expression plasmids containing the CaMV 35S, napin, and chimeric promoters.

Detailed construction details are described in Figs. 7 through to 12.

Initially, a set of constructs were obtained in which genes representing fusion proteins comprised of the TMV MP, His-enriched sequences, and GFP, were placed under the control of the 35S CaMV promoter. His-enriched sequences of three different lengths (namely, those with two, four, and six repeats of the His-encoding sequence unit, example 8) were selected. The MP gene fused to two His-coding units was excised from pMP-2x with *Eco*RI and *Sal*I, the GFP gene was excised from pRT-GFP as a *Xho*I-*Bam*HI-fragment, and both DNA fragments were ligated into pRT101 digested with *Eco*RI and *Bam*HI to give the construct pRT-MP-2x-GFP. Similarly, to obtain pRT-MP-4x-GFP, the MP gene fused to four His-coding units was excised from pMP-4x with *Eco*RI and *Sal*I, the GFP gene was excised from pRT-GFP as a *Xho*I-*Bam*HI-fragment, and both DNA fragments were ligated into pRT101 digested with *Eco*RI and *Bam*HI. To obtain pRT-MP-6x-GFP, the MP gene fused to six His-coding units was excised from pMP-6x with *Eco*RI and *Sal*I, the GFP gene was excised from pRT-GFP as a *Xho*I-*Bam*HI-fragment, and both DNA fragments were ligated into pRT101 digested with *Eco*RI and *Bam*HI.

To obtain equivalent constructs under the control of the napin promoter (replacing the CaMV 35S promoter), the following cloning steps were carried out. The N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K; the fragment containing the rest of the MP gene fused to two His-coding units and the GFP gene was excised from pRT-MP-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP digested with *Nco*I and *Bam*HI to give pNAP-MP-2x-GFP. In a similar fashion, to obtain pNAP-MP-4x-GFP, the N-terminal portion of the 30K TMV MP gene was excised as the *Nco*I-*Hind*III-fragment from pGEM-30K, and the fragment containing the rest of the MP gene fused to four His-coding units and the GFP gene was excised from pRT-MP-4x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP digested with *Nco*I and *Bam*HI. Finally, to construct pNAP-MP-6x-GFP, the N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, and the fragment containing the rest of the MP gene fused to six His-coding units and the GFP gene was excised from pRT-MP-6x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP digested with *Nco*I and *Bam*HI.

Similar cloning was undertaken to obtain constructs in which the MP-His-GFP fusion genes

were placed under the control of the hybrid promoter. The N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, the fragment containing the rest of the MP gene fused to two His-coding units and the GFP gene was excised from pRT-MP-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pHYB digested with *Nco*I and *Bam*HI to give pHYB-MP-2x-GFP. To obtain pHYB-MP-4x-GFP, the N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, the fragment containing the rest of the MP gene fused to four His-coding units and the GFP gene was excised from pRT-MP-4x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pHYB digested with *Nco*I and *Bam*HI. To construct pHYB-MP-6x-GFP, the N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, the fragment containing the rest of the MP gene fused to six His-coding units and the GFP gene was excised from pRT-MP-6x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pHYB digested with *Nco*I and *Bam*HI. Based on the oleosin gene from pOLE4/11, three sets of expression vectors were constructed that had the CaMV 35S, napin, or hybrid promoters, each set containing His-coding regions of three different lengths.

To construct expression vectors based on the CaMV 35S promoter, the following cloning procedures were carried out. The oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing two His-coding sequence units fused to the GFP gene was excised from pHYB-MP-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI resulting in the construct pRT-OLE-2x-GFP. To obtain pRT-OLE-4x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing four His-coding sequence units fused to the GFP gene was excised from pHYB-MP-4x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI. To obtain pRT-OLE-6x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing six His-coding sequence units fused to the GFP gene was excised from pHYB-MP-6x-GFP with *Xho*I and *Bam*HI, and then both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI.

A similar cloning scheme was applied to construct expression vectors based on the napin

promoter. The oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing two His-coding sequence units fused to the GFP gene was excised from pHYB-MP-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pNAP digested with *Nco*I and *Bam*HI, resulting in construct pNAP-OLE-2x-GFP. To obtain pNAP-OLE-4x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing four His-coding sequence units fused to the GFP gene was excised from pHYB-MP-4x-GFP plasmid with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pNAP digested with *Nco*I and *Bam*HI. To obtain pNAP-OLE-6x-GFP, the oleosin gene was excised from pOLE4/11 as the *Nco*I-*Xho*I-fragment, the region containing six His-coding sequence units fused to the GFP gene was excised from pHYB-MP-6x-GFP plasmid with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pNAP digested with *Nco*I and *Bam*HI.

The set of expression vectors based on the hybrid promoter was constructed similarly. The oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing two His-coding sequence units fused to the GFP gene was excised from pHYB-MP-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pHYB digested with *Nco*I and *Bam*HI, resulting in construct pHYB-OLE-2x-GFP. To obtain pHYB-OLE-4x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing four His-coding sequence units fused to the GFP gene was excised from pHYB-MP-4x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pHYB digested with *Nco*I and *Bam*HI. To obtain pHYB-OLE-6x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing six His-coding sequence units fused to the GFP gene was excised from pHYB-MP-6x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pHYB digested with *Nco*I and *Bam*HI.

EXAMPLE 17

Cloning of Cys/Met-enriched oleosin and 30K TMV MP genes into expression plasmids containing the CaMV 35S, napin, and chimeric promoters.

Cys/Met-enriched sequences of three different lengths (with two, four, and six repeats of the

Cys/Met-encoding sequence unit, example 8) were selected. The MP gene fused to two Cys/Met-coding units was excised from pMP-Cys/Met-2x with *Eco*RI and *Sal*I, the GFP gene was excised from pRT-GFP as the *Xho*I-*Bam*HI-fragment, and both DNA fragments were ligated into pRT101 digested with *Eco*RI and *Bam*HI to give the construct pRT-MP-Cys/Met-2x-GFP. The MP gene fused to four Cys/Met-coding units was excised from pMP-Cys/Met-4x with *Eco*RI and *Sal*I, the GFP gene was excised from pRT-GFP as a *Xho*I-*Bam*HI-fragment, and both DNA fragments were ligated into pRT101 digested with *Eco*RI and *Bam*HI to obtain pRT-MP-Cys/Met-4x-GFP.

To obtain equivalent constructs under the control of the napin promoter (replacing the CaMV 35S promoter), the N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K; the fragment containing the rest of the MP gene fused to two Cys/Met-coding units and the GFP gene was excised from pRT-MP-Cys/Met-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP digested with *Nco*I and *Bam*HI to give pNAP-MP-Cys/Met-2x-GFP. In a similar fashion, to obtain pNAP-MP-Cys/Met-4x-GFP, the N-terminal portion of the 30K TMV MP gene was excised as the *Nco*I-*Hind*III-fragment from pGEM-30K, and the fragment containing the rest of the MP gene fused to four Cys/Met-coding units and the GFP gene was excised from pRT-MP-Cys/Met-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP digested with *Nco*I and *Bam*HI. Finally, to construct pNAP-MP-Cys/Met-6x-GFP, the N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, and the fragment containing the rest of the MP gene fused to six Cys/Met-coding units and the GFP gene was excised from pRT-MP-Cys/Met-6x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP digested with *Nco*I and *Bam*HI.

Similar cloning was undertaken to obtain the constructs in which the MP-Cys/Met-GFP fusion genes were placed under the control of the hybrid promoter (HYB). The N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, the fragment containing the rest of the MP gene fused to two Cys/Met-coding units and the GFP gene was excised from pRT-MP-Cys/Met-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pHYB digested with *Nco*I and *Bam*HI to give pHYB-MP-Cys/Met-2x-GFP. To obtain pHYB-MP-Cys/Met-4x-GFP, the N-terminal portion of the

TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, the fragment containing the rest of the MP gene fused to four His-coding units and the GFP gene was excised from pRT-MP-Cys/Met-4x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pHYB digested with *Nco*I and *Bam*HI. To construct pHYB-MP-Cys/Met-6x-GFP, the N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, the fragment containing the rest of the MP gene fused to six Cys/Met-coding units and the GFP gene was excised from pRT-MP-Cys/Met-6x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pHYB digested with *Nco*I and *Bam*HI. Based on the oleosin gene from pOLE4/11, three sets of expression vectors were constructed that had the CaMV 35S, napin, or hybrid promoters, each set containing Cys/Met-coding regions of three different lengths, as in example 16.

For construction of expression vectors based on the CaMV 35S promoter, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing two Cys/Met-coding sequence units fused to the GFP gene was excised from pHYB-MP-Cys/Met-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI resulting in the construct pRT-OLE-Cys/Met-2x-GFP. To obtain pRT-OLE-Cys/Met-4x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing four Cys/Met-coding sequence units fused to the GFP gene was excised from pHYB-MP-Cys/Met-4x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI. To obtain pRT-OLE-Cys/Met-6x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing six Cys/Met-coding sequence units fused to the GFP gene was excised from pHYB-MP-Cys/Met-6x-GFP with *Xho*I and *Bam*HI, and then both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI.

Similarly, for expression vectors based on the napin promoter, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing two Cys/Met-coding sequence units fused to the GFP gene was excised from pHYB-MP-Cys/Met-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pNAP digested with *Nco*I and *Bam*HI, resulting in construct pNAP-OLE-Cys/Met-2x-GFP. To obtain pNAP-OLE-Cys/Met-4x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the

region containing four Cys/Met-coding sequence units fused to the GFP gene was excised from pHYB-MP-Cys/Met-4x-GFP plasmid with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pNAP digested with *Nco*I and *Bam*HI. To obtain pNAP-OLE-Cys/Met-6x-GFP, the oleosin gene was excised from pOLE4/11 as the *Nco*I-*Xho*I-fragment, the region containing six Cys/Met-coding sequence units fused to the GFP gene was excised from pHYB-MP-Cys/Met-6x-GFP plasmid with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pNAP digested with *Nco*I and *Bam*HI.

The set of expression vectors based on the hybrid promoter was constructed in a similar manner. The oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, and the region containing two Cys/Met-coding sequence units fused to the GFP gene was excised from pHYB-MP-Cys/Met-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pHYB digested with *Nco*I and *Bam*HI, resulting in construct pHYB-OLE-Cys/Met-2x-GFP. To obtain pHYB-OLE-Cys/Met-4x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing four Cys/Met-coding sequence units fused to the GFP gene was excised from pHYB-MP-Cys/Met-4x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pHYB digested with *Nco*I and *Bam*HI. To obtain pHYB-OLE-Cys/Met-6x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, and the region containing six Cys/Met-coding sequence units fused to the GFP gene was excised from pHYB-MP-Cys/Met-6x-GFP with *Xho*I and *Bam*HI. Both DNA fragments were ligated into pHYB digested with *Nco*I and *Bam*HI.

EXAMPLE 18

Cloning of glycine-enriched 30K TMV MP and oleosin genes into expression plasmids containing the CaMV 35S, napin, and chimeric promoters.

Glycine-enriched sequences of two different lengths (with two and four repeats of the Gly-encoding sequence units) were selected. The MP gene fused to two Gly-coding units was excised from pMP-Gly-2x with *Eco*RI and *Sal*I, the GFP gene was excised from pRT-GFP as the *Xho*I-*Bam*HI-fragment, and both DNA fragments were ligated into pRT101 digested with *Eco*RI and *Bam*HI to give the construct pRT-MP-Gly-2x-GFP. The MP gene fused to four

Gly-coding units was excised from pMP-Gly-4x with *Eco*RI and *Sal*I, the GFP gene was excised from pRT-GFP as a *Xho*I-*Bam*HI-fragment, and both DNA fragments were ligated into pRT101 digested with *Eco*RI and *Bam*HI to obtain pRT-MP-Gly-4x-GFP.

To obtain equivalent constructs under the control of the napin promoter (replacing the CaMV 35S promoter), the *Nco*I-*Hind*III-fragment of the TMV MP gene was excised from pGEM-30K (the fragment containing the rest of the MP gene fused to two Gly-coding units) and the GFP gene was excised from pRT-MP-Gly-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP digested with *Nco*I and *Bam*HI to give pNAP-MP-Gly-2x-GFP. To obtain pNAP-MP-Gly-4x-GFP, the N-terminal portion of the 30K TMV MP gene was excised as the *Nco*I-*Hind*III-fragment from pGEM-30K, and the fragment containing the rest of the MP gene fused to four Gly-coding units and the GFP gene was excised from pRT-MP-Gly-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP digested *with Nco*I and *Bam*HI.

To obtain the hybrid promoter construct (HYB), the N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, and the GFP gene was excised from pRT-MP-Gly-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pHYB digested with *Nco*I and *Bam*HI to give pHYB-MP-Gly-2x-GFP. To obtain pHYB-MP-Gly-4x-GFP, the N-terminal portion of the TMV MP gene was excised as a *Nco*I-*Hind*III-fragment from pGEM-30K, the fragment containing the rest of the MP gene fused to two glycine-coding units and the GFP gene was excised from pRT-MP-Gly-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pHYB digested with *Nco*I and *Bam*HI.

Based on the oleosin gene from pOLE4/11, three sets of expression vectors were constructed that had the CaMV 35S, napin, or hybrid promoters, each set containing Gly-coding regions of two different lengths, as in example 16. For construction of expression vectors based on the CaMV 35S promoter, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing two Gly-coding sequence units fused to the GFP gene was excised from pHYB-MP-Gly-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI resulting in the construct pRT-OLE-Gly-2x-GFP. To obtain pRT-OLE-Gly-4x-GFP, the oleosin gene was excised from pOLE4/11 as

a *Nco*I-*Xho*I-fragment, the region containing four Gly-coding sequence units fused to the GFP gene was excised from pHYB-MP-Gly-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI.

Similarly, for expression vectors based on the napin promoter, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing two Gly-coding sequence units fused to the GFP gene was excised from pHYB-MP-Gly-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pNAP digested with *Nco*I and *Bam*HI, resulting in construct pNAP-OLE-Gly-2x-GFP. To obtain pNAP-OLE-Gly-4x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing four Gly-coding sequence units fused to the GFP gene was excised from pHYB-MP-Gly-4x-GFP plasmid with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pNAP digested with *Nco*I and *Bam*HI.

The set of expression vectors based on the hybrid promoter was constructed in a similar manner. The oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, and the region containing two Gly-coding sequence units fused to the GFP gene was excised from pHYB-MP-Gly-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pHYB digested with *Nco*I and *Bam*HI, resulting in construct pHYB-OLE-Gly-2x-GFP. To obtain pHYB-OLE-Gly-4x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing four Gly-coding sequence units fused to the GFP gene was excised from pHYB-MP-Gly-4x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pHYB digested with *Nco*I and *Bam*HI.

EXAMPLE 19

Cloning of lysine-enriched 30K TMV MP and oleosin genes into expression plasmids containing the CaMV 35S, napin, and chimeric promoters.

Lysine-enriched sequences of two different lengths (with two and four repeats of the Lys-encoding sequence units) were selected. The MP gene fused to two Lys-coding units was excised from pMP-Lys-2x with *Eco*RI and *Sal*I, the GFP gene was excised from pRT-GFP as

the *Xho*I-*Bam*HI-fragment, and both DNA fragments were ligated into pRT101 digested with *Eco*RI and *Bam*HI to give the construct pRT-MP-Lys-2x-GFP. The MP gene fused to four Lys-coding units was excised from pMP-Lys-4x with *Eco*RI and *Sal*I, the GFP gene was excised from pRT-GFP as a *Xho*I-*Bam*HI-fragment, and both DNA fragments were ligated into pRT101 digested with *Eco*RI and *Bam*HI to obtain pRT-MP-Lys-4x-GFP, as in examples 16-18.

To obtain equivalent constructs under the control of the napin and hybrid (HYB) promoters (replacing the CaMV 35S promoter), *Nco*I-*Hind*III-fragment of TMV MP gene was excised from pGEM-30K (the fragment containing the rest of the MP gene fused to two Lys-coding units) and the GFP gene was excised from pRT-MP-Lys-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP or pHYB digested with *Nco*I and *Bam*HI to give pNAP-MP-Lys-2x-GFP and pHYB-MP-Lys-2x-GFP, respectively.

To obtain pNAP-MP-Lys-4x-GFP and pHYB-MP-Lys-4x-GFP the N-terminal portion of the 30K TMV MP gene was excised as the *Nco*I-*Hind*III-fragment from pGEM-30K, and the fragment containing the rest of the MP gene fused to four Lys-coding units and the GFP gene was excised from pRT-MP-Lys-2x-GFP with *Hind*III and *Xba*I, and both fragments were ligated into pNAP and pHYB-digested with *Nco*I and *Bam*HI, respectively.

Based on the oleosin gene from pOLE4/11, three sets of expression vectors were constructed that had the CaMV 35S, napin, or hybrid promoters, each set containing Lys-coding regions of two different lengths, as in example 18.

For construction of expression vectors based on the CaMV 35S, napin or HYB promoters, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing two Lys-coding sequence units fused to the GFP gene was excised from pHYB-MP-Lys-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI resulting in the constructs pRT-OLE-Lys-2x-GFP. pNAP-OLE-Lys-2x-GFP and pHYB-OLE-Lys-2x-GFP. To obtain pRT-OLE-Lys-4x-GFP, pNAP-OLE-Lys-4x-GFP, and pHYB-OLE-Lys-4x-GFP, the oleosin gene was excised from pOLE4/11 as a *Nco*I-*Xho*I-fragment, the region containing four Lys-coding sequence units fused to the GFP

gene was excised from pHYB-MP-Lys-2x-GFP with *Xho*I and *Bam*HI, and both DNA fragments were ligated into pRT100 digested with *Nco*I and *Bam*HI.

EXAMPLE 20

Testing the activity of napin and chimeric promoters using transient expression of the GFP gene in plant cells transformed by particle bombardment

To analyze relative transcription levels provided by the napin promoter in pNAP, the hybrid promoter in pHYB, and the CaMV 35S promoter in pRT100, the coding region of the reporter-protein, GFP, was cloned into these plasmids. For construction of GFP-containing clones, a gene for a red-shifted GFP mutant (S65T) was used. The GFP coding region was cloned as a *Nco*I-*Bam*HI-fragment into similarly digested pRT100, pNAP and pHYB.

Epidermal cells transiently transformed using particle bombardment (example 5.) with pNAP-GFP and pHYB-GFP (containing the GFP coding region under the control of the napin and chimeric promoters, respectively) expressed GFP brightly. Observed GFP expression levels were comparable to the control, pRT-GFP, in which GFP expression was under the control of the CaMV 35S promoter. These experiments indicated that both pNAP-GFP and pHYB-GFP expression cassettes are fully functional in living plant cells.

EXAMPLE 21

Cloning of expression cassettes into binary vectors and transformation of corresponding binary vectors into *Agrobacterium*

Gene cassettes containing GFP-tagged histidine, cysteine/methionine, glycine and lysine-enriched oleosin and TMV MP genes under control of CaMV 35S, napin and chimeric promoters from pRT-derivatives were cloned into the T-DNA-located *Hind*III site of the plant transformation binary vector pBin19 prior to *Agrobacterium* transformation.

Agrobacterium tumefaciens AGLI carrying the hyper-virulent, attenuated tumor-inducing plasmid pTi Bo542 (Lazo et al., 1991. Biotechnology 3, 963-967) were transformed by the freeze-thaw method of Holsters (Holsters et al., 1978. Mol.Gen.Genet. 163, 181-187) with some modifications. *Agrobacteria* were cultivated on LB-agar plates supplemented with 0,5 g l^{-1} $MgSO_4$ for two days at 26°C, followed by cultivation on liquid 2X YT medium supplemented with 50 µg/ml rifampicin to mid-exponential phase, then washed with 150 mM NaCl and resuspended in 20 mM $CaCl_2$ at ~1010 cfu ml^{-1} prior to transformation. Aliquots of cell suspension were frozen in liquid nitrogen after addition of glycerol to a final concentration of 15%. After the addition of 5 µg of DNA to cells, the cells were thawed and incubated on ice for 15 minutes, prior to heat-shock at 37°C for 5 minutes. Cells were diluted with fresh 2X YT medium and cultured for 2 hours at 28°C with vigorous aeration prior to being transferred to 1 % agar plates containing 50g µg ml^{-1} rifampicin and 100µg ml^{-1} kanamycin. Kanamycin-resistant colonies were screened by PCR-analysis for the presence/absence of recombinant constructs.

EXAMPLE 22

Agrobacterium-mediated transformation of *Brassica campestris*

Transformation of *Brassica campestris* plants was done essentially as described (Moloney et al., Plant Cell Reports 8, 238-272, 1989). Cotyledons were co-cultivated with appropriate *Agrobacterium* cell suspensions for 2-3 days on 1 % agar with Medium I (Murashige Minimal Organics (MMO) with 3% sucrose, 4 mg/l BAP, 0.7% Phytoagar) at 22°C. Cotyledons were rinsed three times in distilled water prior to transfer to petri dishes containing Medium II (MMO with 4 mg/L BAP, 3% sucrose, 300 mg/l Ticarcillin (DUCHEFA) and 0.7% Phytoagar). After a further 7-10 days, explants were transferred to Medium III (MMO with 4 mg/L BAP, 3% sucrose, 300 mg/l Ticarcillin, 25 mg/l Kanamycin and 0.7% Phytoagar). After additional 2-3 weeks incubation explants with green calli or immature green shoots were transferred to fresh Medium III. For shoot formation, calli and immature green shoots were transfered to Medium IV (MMO with 3% sucrose, 300 mg/l Ticarcillin, 25 mg/l Kanamycin and 0.7% Phytoagar). Fully formed shoots were transferred

to Medium V (MMO with 0.2 mg/l IBA, 3% sucrose, 300 mg/l Ticarcillin and 0.7% Phytoagar) for rooting. Once an established root system had formed, shoots with roots were removed from agar and transferred to moist potting soil, and grown in 16h light/8h dark photoperiod at light intensity of 12.1μmol m^{-2} s^{-1} at 22°C.

EXAMPLE 23

Expression of constructs with various lengths of histidine, cysteine/methionine, glycine and lysine in tobacco epidermal cells

The length of the engineered poly-amino acid sequences were thought to be crucial for the expression of the His, Cys/Met, Gly and Lys-constructs, for two reasons:

1) Stability of the long polyamino acid sequences (for example) in the 8xHis-series were unstable in an *in vitro* translation system and the 6xHis-series restricted the growth of *E.coli* in liquid medium (example 8.).
2) The inherent limitation of cells to produce large amounts of polyamino acids without having an effect on other proteins and their synthesis.

Therefore, all CaMV 35S promoter-driven constructs were tested *in planta* to examine the limitation(s) of the large polyamino acid sequences in plant cells. Despite the suspected limitation(s), all constructs (as prepared in examples 10-13) were expressed in tobacco epidermal cells. However, constructs containing longer polyamino acid sequences were relatively less well expressed than shorter polyamino acid sequences, therefore, constructs with four His, Cys/Met, Gly and Lys-cassettes were chosen for stable transformation experiments.

EXAMPLE 24

The *in planta* production of oleosin-histidine and MP-histidine fusion proteins

Western blot analysis of several plant lines (Table 1.) showed that antibodies raised against

histidine sequences, MP, oleosin or GFP (Figs. 13.-15.) could be used to detect proteins with known molecular weights (MWs). The predicted MWs of oleosin, MP, GFP and 4X His-cassette (with 14 His-codons in each cassette) were 18.5, 30, 30, and 2.5 kDa, respectively. The appropriate MWs of fusion proteins can be calculated from these values.

These fusion proteins were shown to be identical in size to the predicted molecular weight, when examined with SDS-PAGE gel electrophoresis and by Western blot analysis. Stability of the transgene-encoded protein products (oleosin-His and MP-His) was investigated by Western blot analysis. Analysis of (selfed) plant generations (3rd generation) after transformation (Table 1) showed clearly that transgene expression was stable over successive generations. Moreover, Western blot analysis of examined transformed plants further revealed that the size of protein product remained constant regardless of the plant generation. These analyses also showed that the amount of fusion protein product was relatively constant between plant generations indicative of an acquired stable expression level relative to other plant proteins.

Table 1. Response of transgenic *Brassica campestris* plant lines to specific antibodies generated against the protein products encoded by the appropriate transgenes.

Transgenic Plant line	**Fusion protein**	**Antibody response**		
		Oleosin	**His**	**MP**
5.1A11	MP	ND	+++	+++
5.1A18	MP	ND	+++	+++
8.26	MP	ND	+	+
8.29	MP	ND	+	+
14.17C2	Oleosin	++	++	ND
14.5C8	Oleosin	+	+	ND
14.12A10	Oleosin	+++	+++	ND

ND = not detected, + = weak response, ++ = good response and +++ = strong response. Transgenes in lines 5/14 and 8/17 are encoded by NAP-promoter and HYB-promoter, respectively.

EXAMPLE 25

Amino acid analysis of seeds from transgenic *Brassica campestris* plant lines expressing four histidine cassettes

Replicate samples of seeds destined for amino acid analysis were obtained from transgenic *Brassica* lines expressing four histidine cassettes (see Figure 3). Amino acid analyses were carried out on the total protein sample obtained from seeds (Table 2).

Table 2. Seeds used for amino acid analysis.

Plant line	Amount of seeds	Seed weight (mg)	Protein weight (mg)
WT	10	33	11.3
14.5C8	8	11.5	7.2
14.12A10	9	30.9	19.4

Transgenic *Brassica campestris* plant line 14.12A10 contained 30 % more histidine compared to wild-type plant (Table 3.). The increased amount of histidine did not accumulate at the expense of other amino acids, but altered the amount of total proteins.

Table 3. Relative increase of the histidine in transgenic *B. campestris* seeds.

Plant line	Amount of histidine (g) per protein (kg)	Increase of histidine %
WT	3,20	0
14.5C8	3.20	0
14.12A10	4.74	32

L5

Claims:

1. A method for increasing the content of a selected amino acid in plant tissue, **characterized** in that the method comprises the steps of transforming any crop plant or any other plant species with at least one recombinant nucleotide sequence construct comprising tissue and/or organ specific promoter(s) driving transcription during selected stages of morphogenesis operably linked to a chimeric nucleotide sequence comprising;
(i) a nucleotide sequence lacking a termination codon and encoding a carrier-protein selected from plant proteins having a function(s) in the secretory intracellular trafficking pathways enabling accumulation in selected plant tissues; said nucleotide sequence being fused in frame with
(ii) a nucleotide sequence comprising an optimal number of any desired amino acid codons, the number of which enables problem-free protein translation; said construct enabling a targeted accumulation of the carrier-protein with a stably attached polyamino acid extension and biological functions, which are intact relative to the native carrier protein.

2. The method according to claim 1, **characterized** in that the construct enabling a targeted accumulation of the carrier protein with intact biological functions and a stable polyamino acid extension is obtainable by a method comprising the steps
(a) fusing the recombinant nucleotide sequence construct of claim 1 in-frame with a nucleotide sequence encoding a reporter-protein;
(b) identifying and selecting in a cell free translation system constructs in which the number of amino acid codons can be translated without problems;
(c) selecting with a transient expression assay combined with micro-projectile particle bombardment of crop plant embryos or leaves and confocal laser scanning microscopy, constructs enabling targeted accumulation of the carrier-protein with a stable polyamino acid extension by detecting the reporter-protein in living plant cells;
(d) providing *Agrobacterium* strains transformed with constructs selected in step (c) and selecting positive clones using Southern blot analysis;
(e) transforming transgenic crop plants with constructs selected in step (d) and verifying the selecting clones carrying the correct construct with PCR-primers specific for nucleotide

sequence of the selected amino acid codon-containing carrier-protein, Western blotting with antibodies to the carrier protein and detection of reporter-protein in plant cells;
(f) removing the nucleotide sequence encoding the reporter-protein from constructs in the clones selected in step (e).

3. The method according to claim 1, **characterized** in that the polyamino acid sequence is selected from a group consisting of histidine, cysteine, methionine, glycine, lysine, tryptophan, alanine, valine, leucine, isoleucine, proline, phenylalanine, tyrosine, serine, threonine, arginine, aspartate, glutamate, asparagine and glutamine.

4. The method according to claim 1, **characterized** in that the polyamino acid sequence comprises histidine.

5. The method according to claim 1, **characterized** in that the optimal number of amino acid codons in the construct enabling expression of a stable amino acid enriched protein is from four to eighty.

6. The method according to claim 1, **characterized** in that the selected plant tissue is a seed.

7. The method according to claim 1, **characterized** in that the selected plant tissue is cell walls or cell membranes.

8. The method according to claim 1, **characterized** in that the carrier-protein is a cell wall protein or a plant viral movement protein.

9. The method according to claim 1, **characterized** in that the carrier-protein is selected from a group consisting of oleosin, caleosin, steroleosin, cruciferin, napin or a plant viral movement protein.

10. The method according to claim 1, **characterized** in that the carrier-protein is oleosin (OLE).

11. The method according to claim 1, **characterized** in that the carrier-protein is the 30K movement protein of tobacco mosaic virus (TMV MP).

12. The method according to claim 1, **characterized** in that the regulatory sequence is a promoter expressing during embryogenesis.

13. The method according to claim 1, **characterized** in that the regulatory sequence is selected from a group consisting of napin (NAP), CaMV 35S, hybrid promoter (HYB) (comprising CaMV 35S enhancer element and a full napin promoter), CaMV 19S, nopalin, phaseolin, steroleosin, caleosin, cruciferin, AMV, heat-shock, albumin 2S and oleosin.

14. The method according to claim 1, **characterized** in that the regulatory sequence is selected from a group comprising a napin (NAP) promoter, a CaMV 35S promoter and a chimeric hybrid (HYB) promoter.

15. The method according to claim 14, **characterized** in that the napin promoter is a napin (NAP) promoter of *Arabidopsis thaliana*

16. The method according to claim 14, **characterized** in that the 35S promoter is a 35S promoter of CaMV.

17. The method according to claim 14, **characterized** in that the chimeric hybrid (HYB) promoter comprises an enhancer sequence of the 35S promoter and the entire NAP promoter of *Arabidopsis thaliana.*

18. The method according to claim 1, **characterized** in that for targeted accumulation of MP protein in seed the MP protein is expressed under the napin promoter.

19. The method according to claim 2, **characterized** in that the optional nucleotide sequence encoding the reporter-protein is a nucleotide sequence encoding a protein selected from the group consisting of green fluorescent protein (GFP) or its derivatives, any other fluorescing protein such as red fluorescent protein, ß-glucuronidase, obelin and luciferase.

20. A recombinant nucleotide sequence construct for increasing the amino acid content in desired crop plant tissues, **characterized** in that the construct enables targeted accumulation of stable amino acid-enriched protein in plant tissues and cells, said construct comprises tissue or organ specific regulatory sequences driving transcription during selected stages of morphogenesis; operably linked to a chimeric nucleotide sequence comprising;
(a) a nucleotide sequence encoding a protein-carrier with intact biological functions and lacking a termination codon; entailed with
(b) a nucleotide sequence comprising an optimal number of codons encoding the selected, desired amino acid and lacking a termination codon.

21. A recombinant nucleotide sequence construct for obtaining constructs for increasing the amino acid content in crop plants, **characterized** in that said construct enables targeted accumulation of stable amino acid-enriched protein in selected plant tissues comprises regulatory sequences operably linked to a chimeric nucleotide sequence comprising;
(a) a nucleotide sequence encoding a carrier-protein with intact biological functions and lacking a termination codon; entailed with
(b) a nucleotide sequence comprising an optimal number of codons encoding the selected, desired amino acid and lacking a termination codon;
said chimeric nucleotide sequence being fused in transcriptional frame with a nucleotide sequence encoding a reporter-protein.

22. A composition with an increased amino acid content obtained by the method according to claim 1, **characterized** in that the composition comprises accumulated in the targeted plant tissue of an amino acid-enriched protein and compatible additives.

23. The composition according to claim 22, **characterized** in that the amino acid-enriched protein is accumulated in the membranes of membraneous oil bodies or in cell walls of seed.

24. A method for producing a composition according to claim 23, **characterized** in that the amino acid-enriched protein accumulated in the membranes of the membraneous oil bodies or in the cell walls of seeds is recovered from the residues obtained after recovering the oil.

25. Use of the composition according to claim 17 as animal feed, including the use as a direct source of nutrients for human beings.

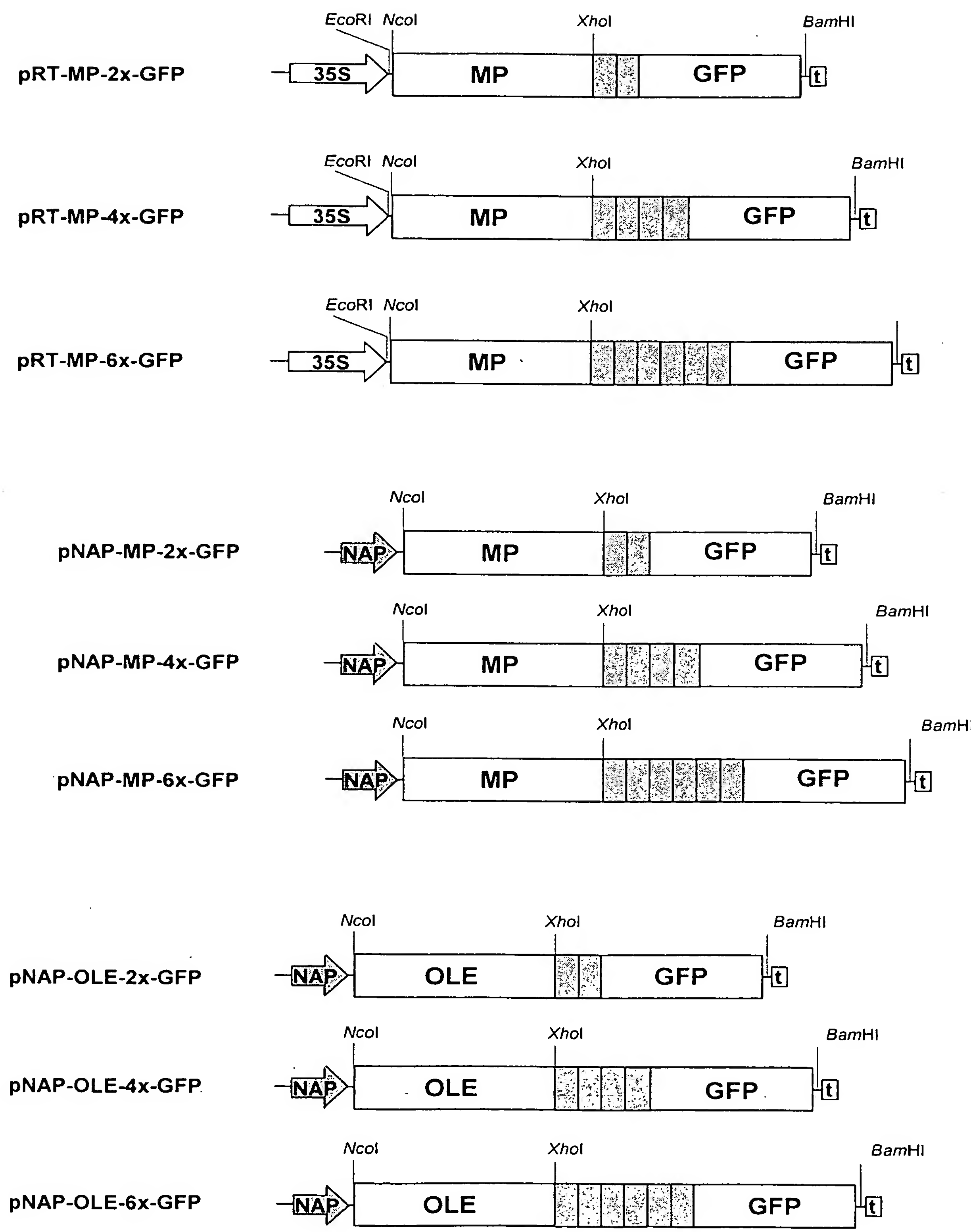
L6

Abstract

The present invention provides a construct and a method for increasing the content of any desired amino acid, particularly histidine, in plant tissue by stable transformation of plants with a construct comprising regulatory sequences operably linked to a chimeric nucleotide sequence wherein a nucleotide sequence coding for a protein-carrier extended by a nucleotide sequence enriched with codons encoding desired amino acids fused in-frame with a nucleotide sequence encoding a reporter-protein. The increased amino acid content in plant tissues, particularly membraneous oil bodies and cell walls of seed provides a composition useful as animal feed.

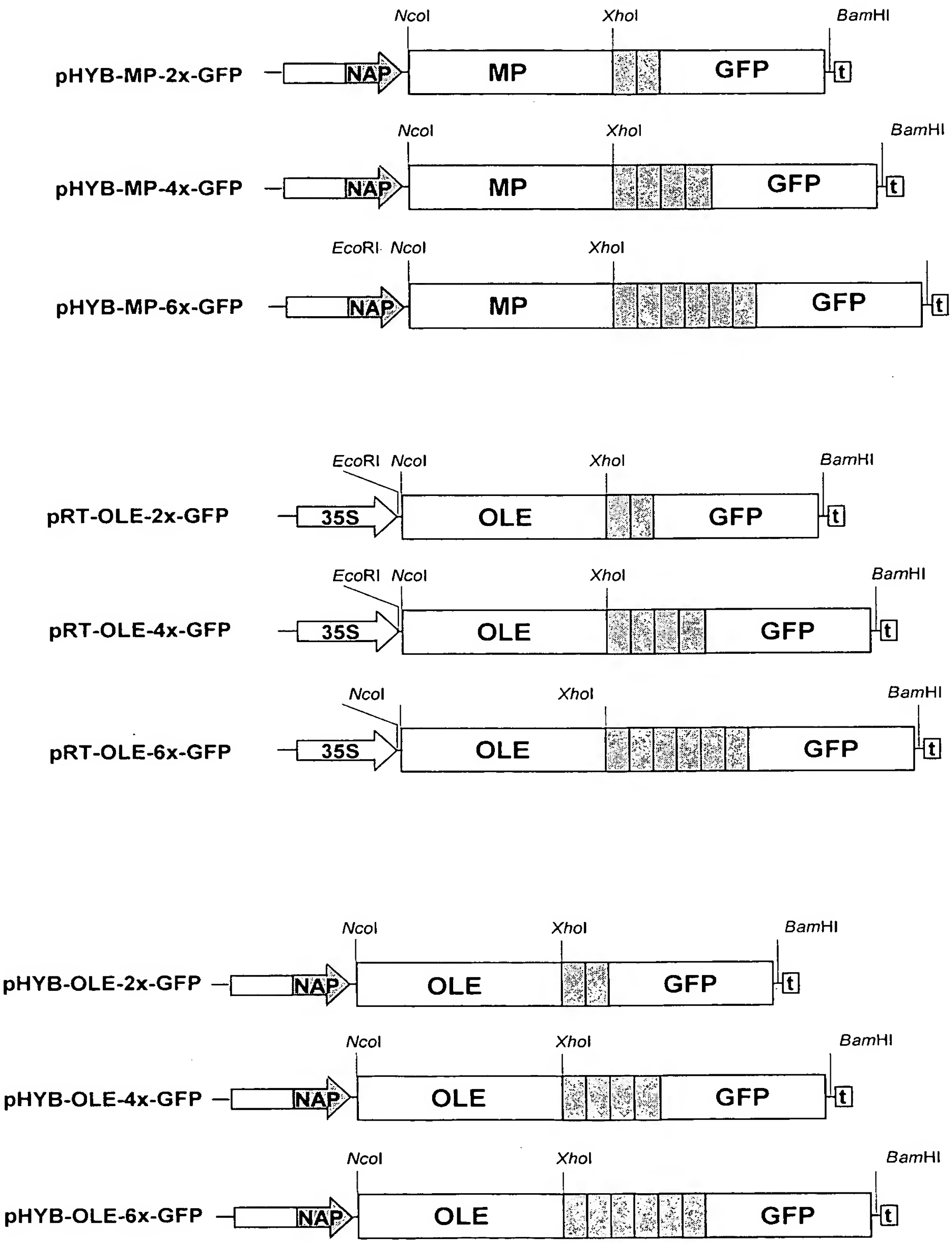
L7

Fig. 1


pRT-MP-2x-GFP
EcoRI NcoI
XhoI
BamHI
35S
MP
GFP
t
pRT-MP-4x-GFP
EcoRI NcoI
XhoI
BamHI
35S
MP
GFP
t
pRT-MP-6x-GFP
EcoRI NcoI
XhoI
35S
MP
GFP
t
pNAP-MP-2x-GFP
NcoI
XhoI
BamHI
NAP
MP
GFP
t
pNAP-MP-4x-GFP
NcoI
XhoI
BamHI
NAP
MP
GFP
t
pNAP-MP-6x-GFP
NcoI
XhoI
BamHI
NAP
MP
GFP
t
pNAP-OLE-2x-GFP
NcoI
XhoI
BamHI
NAP
OLE
GFP
t
pNAP-OLE-4x-GFP
NcoI
XhoI
BamHI
NAP
OLE
GFP
t
pNAP-OLE-6x-GFP
NcoI
XhoI
BamHI
NAP
OLE
GFP
t

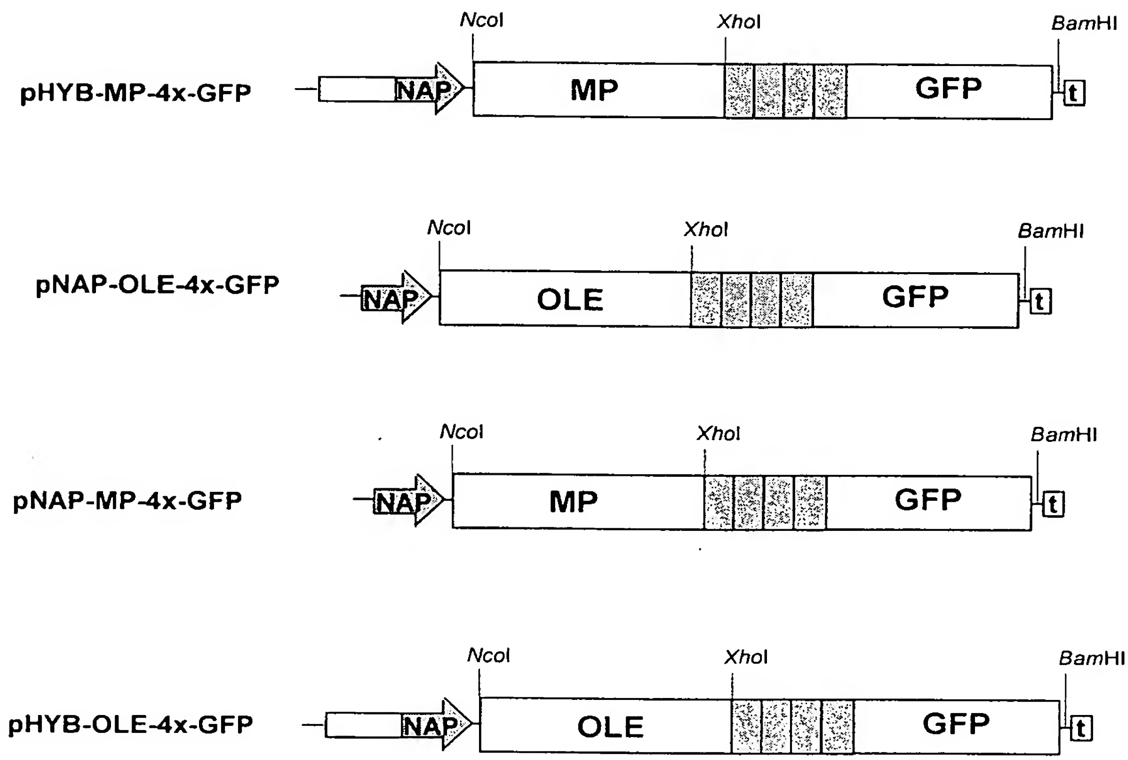
L7

Fig. 1 continued



L7

Fig. 2


pHYB-MP-4x-GFP
NcoI
XhoI
BamHI
NAP
MP
GFP
t
pNAP-OLE-4x-GFP
NcoI
XhoI
BamHI
NAP
OLE
GFP
t
pNAP-MP-4x-GFP
NcoI
XhoI
BamHI
NAP
MP
GFP
t
pHYB-OLE-4x-GFP
NcoI
XhoI
BamHI
NAP
OLE
GFP
t

L7

Fig. 3

A.

*Xho*I *Sal*I *Bam*HI

CTCGAGTTCACCATCACCATCTCCATCACGGGCACCATCACCATCACCATCACCATG**GTCGAC**TTTA**GGATCC**

R V H H H H L H H G H H H H H H H H G R L *

B.

*Xho*I *Sal*I

CAC**CTCGAG**TATGTTGTTGCATGTGCATGTGCTGTTGCAT**GTCGAC**AAAC

R V C C C M C M C C C M S

C.

*Xho*I *Sal*I *Bam*HI

GCGC**CTCGAG**TTGGTGGAGGTGGAGGCGGTGGAGGTGGC**GTCGAC**AAAT**GGATCC**CC

R V G G G G G G G G G V

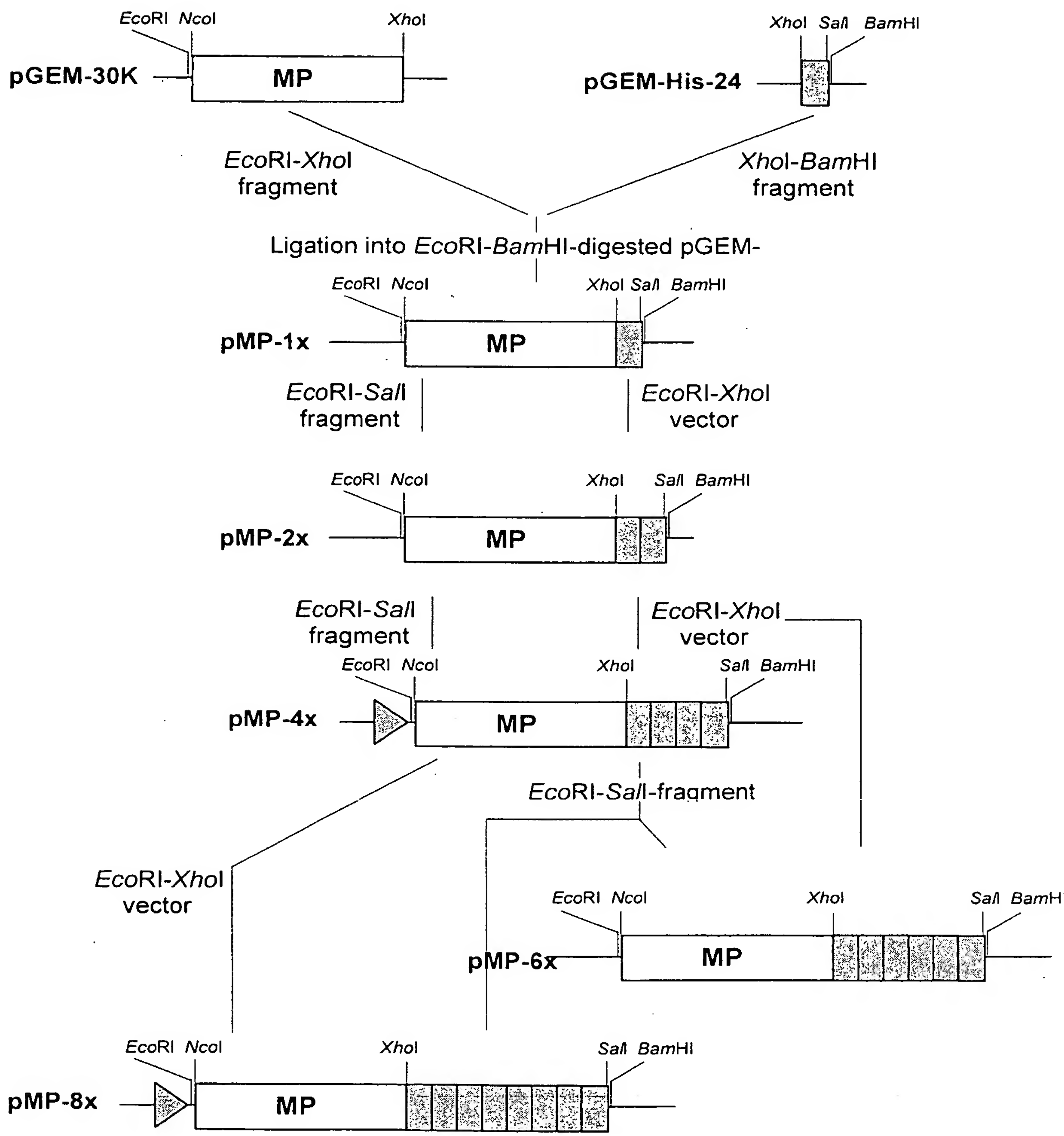
D.

*Xho*I *Sal*I *Bam*HI

GCGC**CTCGAG**TTAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG**GTCGAC**AAAT**GGATCC**CC

R V K K K K K K K K K K K K V

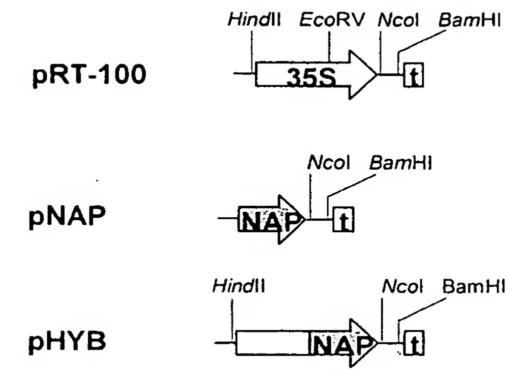
Fig. 4


EcoRI NcoI
XhoI
pGEM-30K
MP
XhoI SalI BamHI
pGEM-His-24
EcoRI-XhoI
fragment
XhoI-BamHI
fragment
Ligation into EcoRI-BamHI-digested pGEM-
EcoRI NcoI
XhoI SalI BamHI
pMP-1x
MP
EcoRI-SalI
fragment
EcoRI-XhoI
vector
EcoRI NcoI
XhoI
SalI BamHI
pMP-2x
MP
EcoRI-SalI
fragment
EcoRI-XhoI
vector
EcoRI NcoI
XhoI
SalI BamHI
pMP-4x
MP
EcoRI-SalI-fragment
EcoRI-XhoI
vector
EcoRI NcoI
XhoI
SalI BamHI
pMP-6x
MP
EcoRI NcoI
XhoI
SalI BamHI
pMP-8x
MP

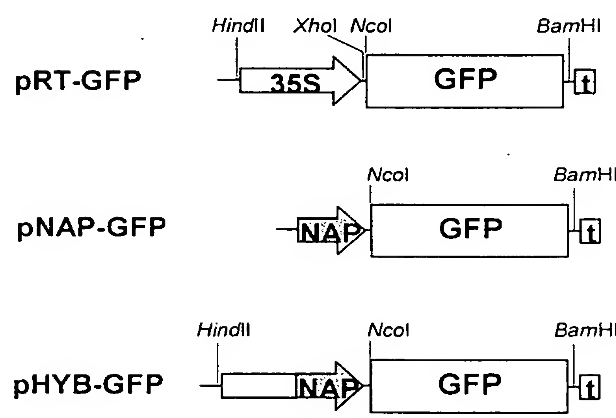
L7

Fig. 5

A.

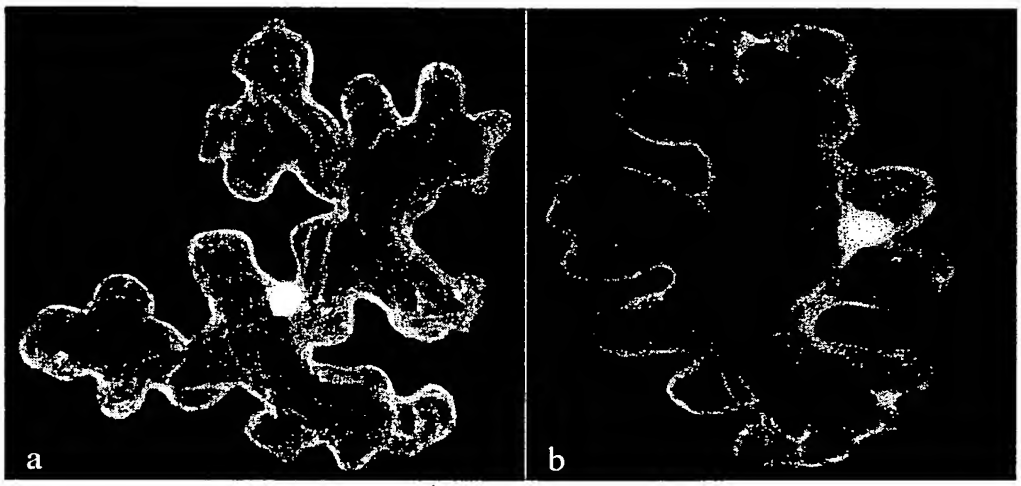

pRT-100
HindII
EcoRV
NcoI
BamHI
35S
t
pNAP
NcoI
BamHI
NAP
t
pHYB
HindII
NcoI
BamHI
NAP
t

B.


pRT-GFP
HindII
XhoI
NcoI
BamHI
35S
GFP
t
pNAP-GFP
NcoI
BamHI
NAP
GFP
t
pHYB-GFP
HindII
NcoI
BamHI
NAP
GFP
t

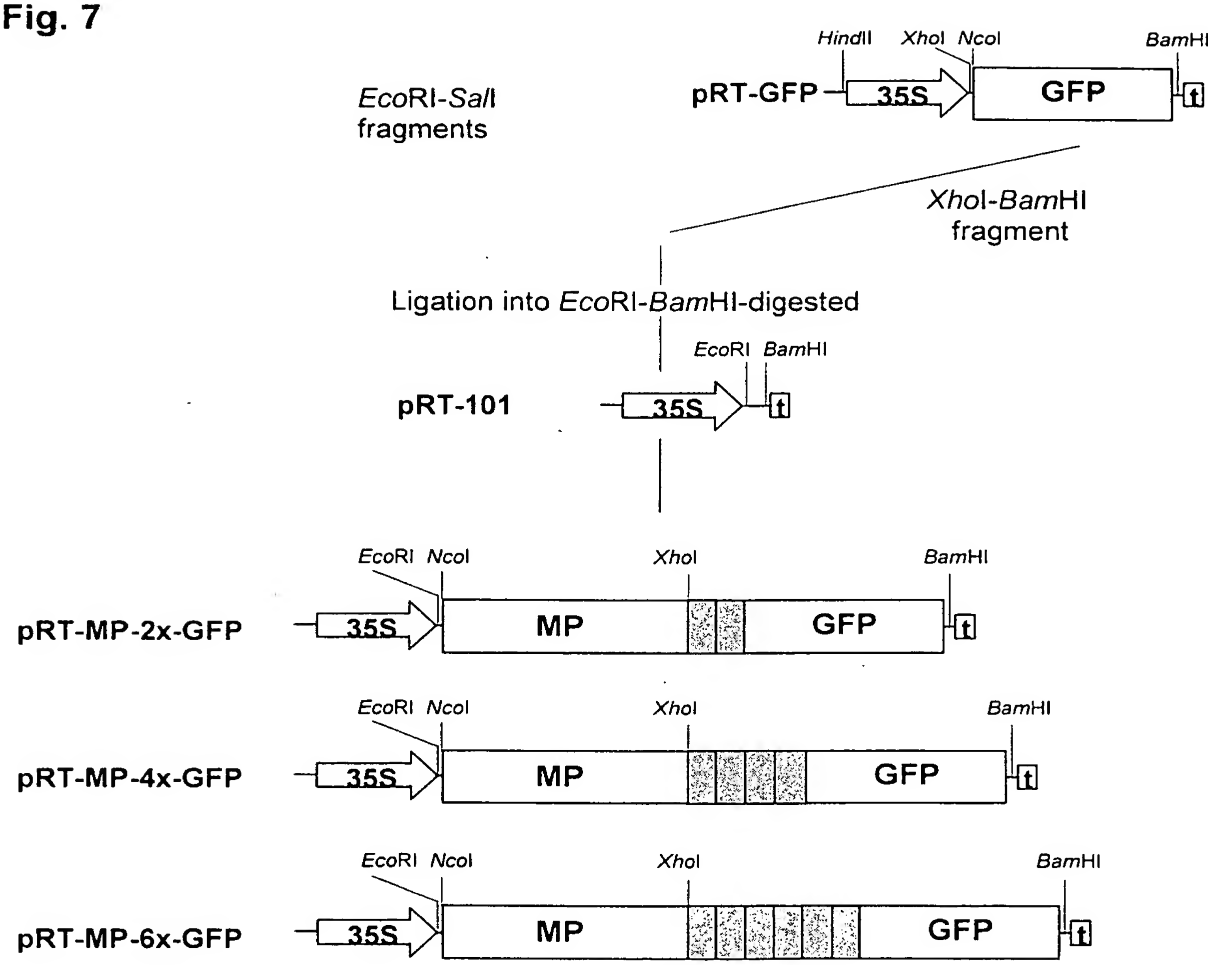
L7

Fig. 6



L 7

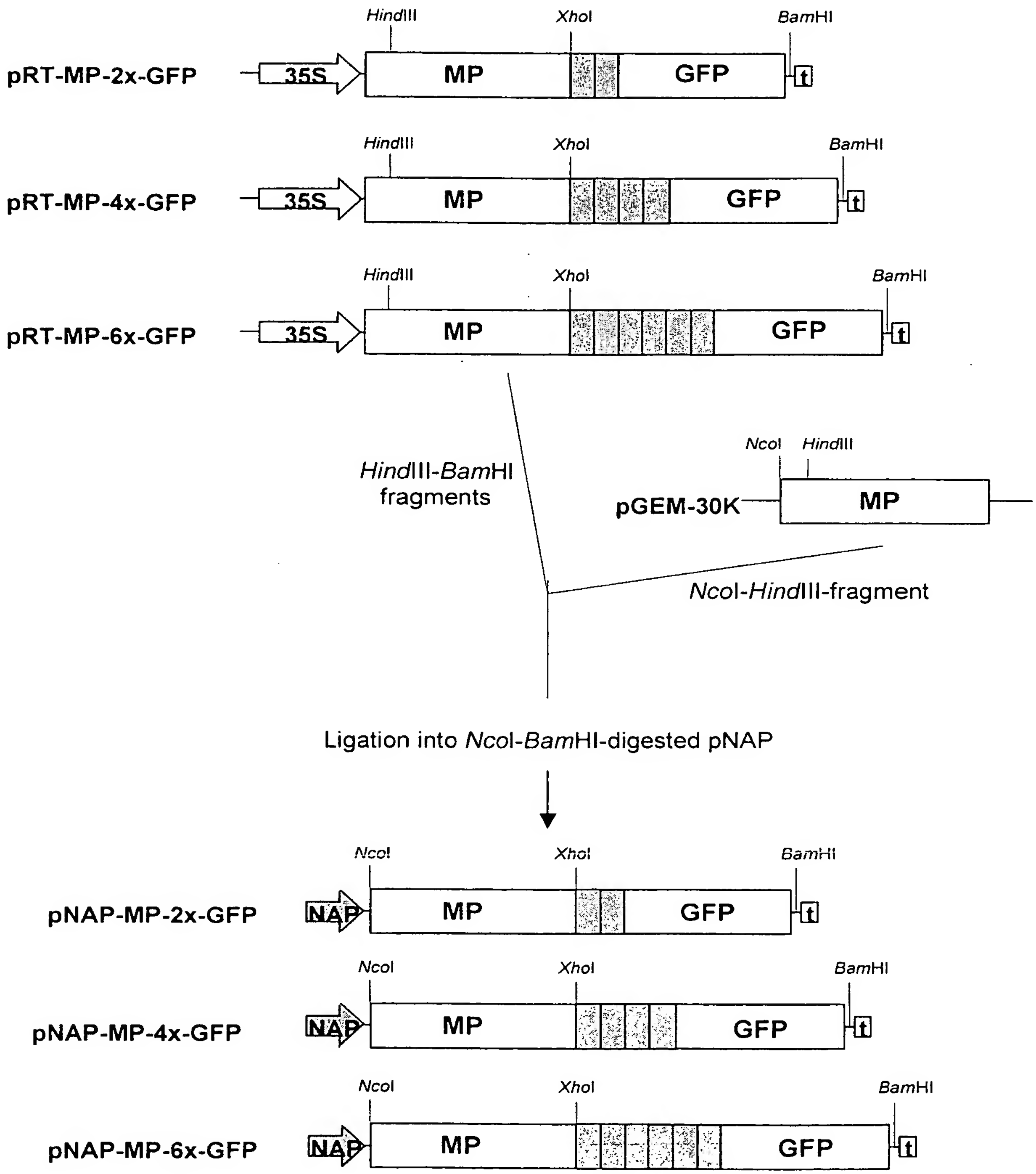
Fig. 7



HindII
XhoI
NcoI
BamHI
pRT-GFP
35S
GFP
t
EcoRI-SalI fragments
XhoI-BamHI fragment
Ligation into EcoRI-BamHI-digested
EcoRI
BamHI
pRT-101
35S
t
EcoRI
NcoI
XhoI
BamHI
pRT-MP-2x-GFP
35S
MP
GFP
t
EcoRI
NcoI
XhoI
BamHI
pRT-MP-4x-GFP
35S
MP
GFP
t
EcoRI
NcoI
XhoI
BamHI
pRT-MP-6x-GFP
35S
MP
GFP
t

L7

Fig. 8


HindIII
XhoI
BamHI
pRT-MP-2x-GFP
35S
MP
GFP
t
pRT-MP-4x-GFP
pRT-MP-6x-GFP
HindIII-BamHI fragments
NcoI
pGEM-30K
NcoI-HindIII-fragment
Ligation into NcoI-BamHI-digested pNAP
pNAP-MP-2x-GFP
NAP
pNAP-MP-4x-GFP
pNAP-MP-6x-GFP

i7

Fig. 9

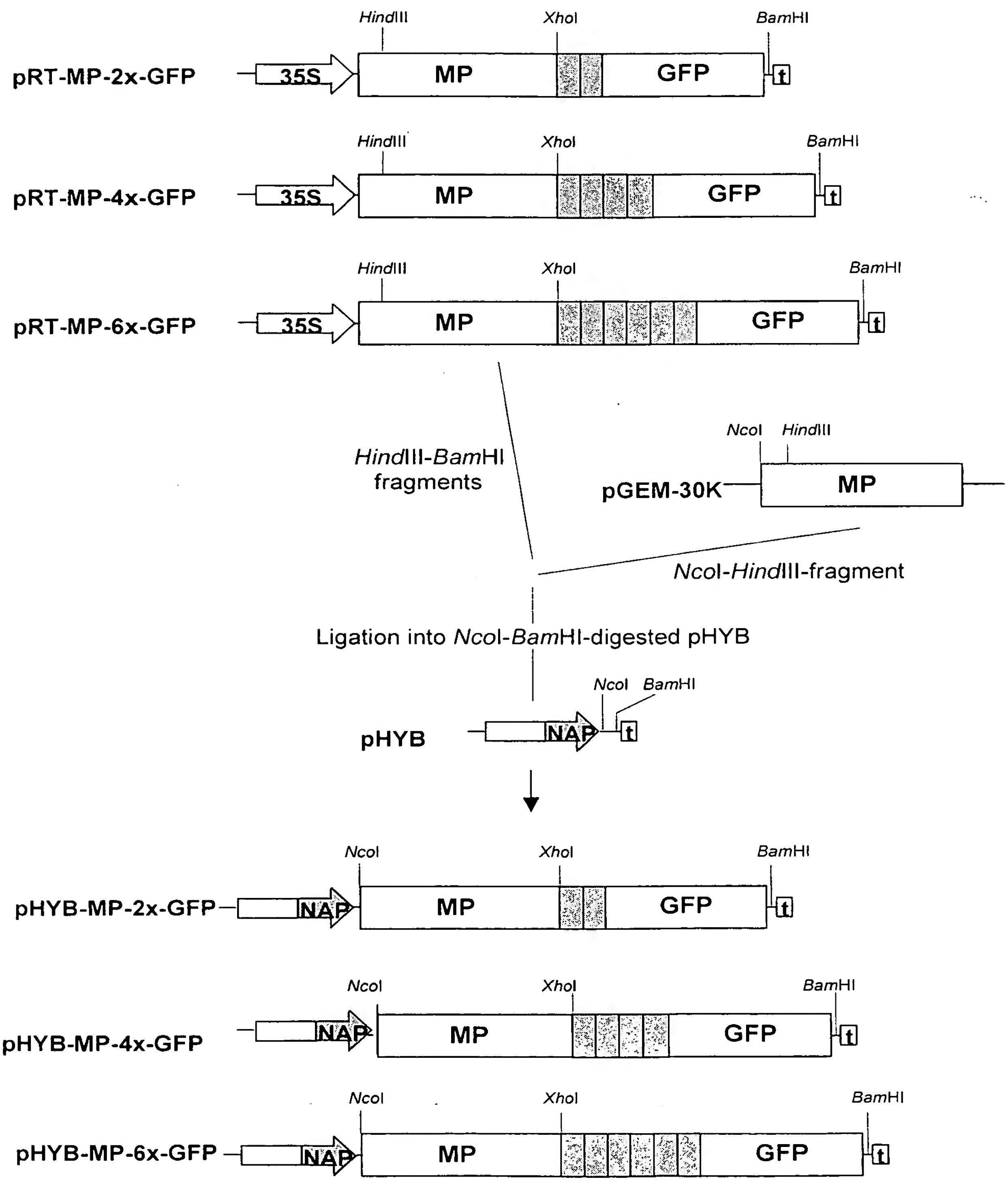
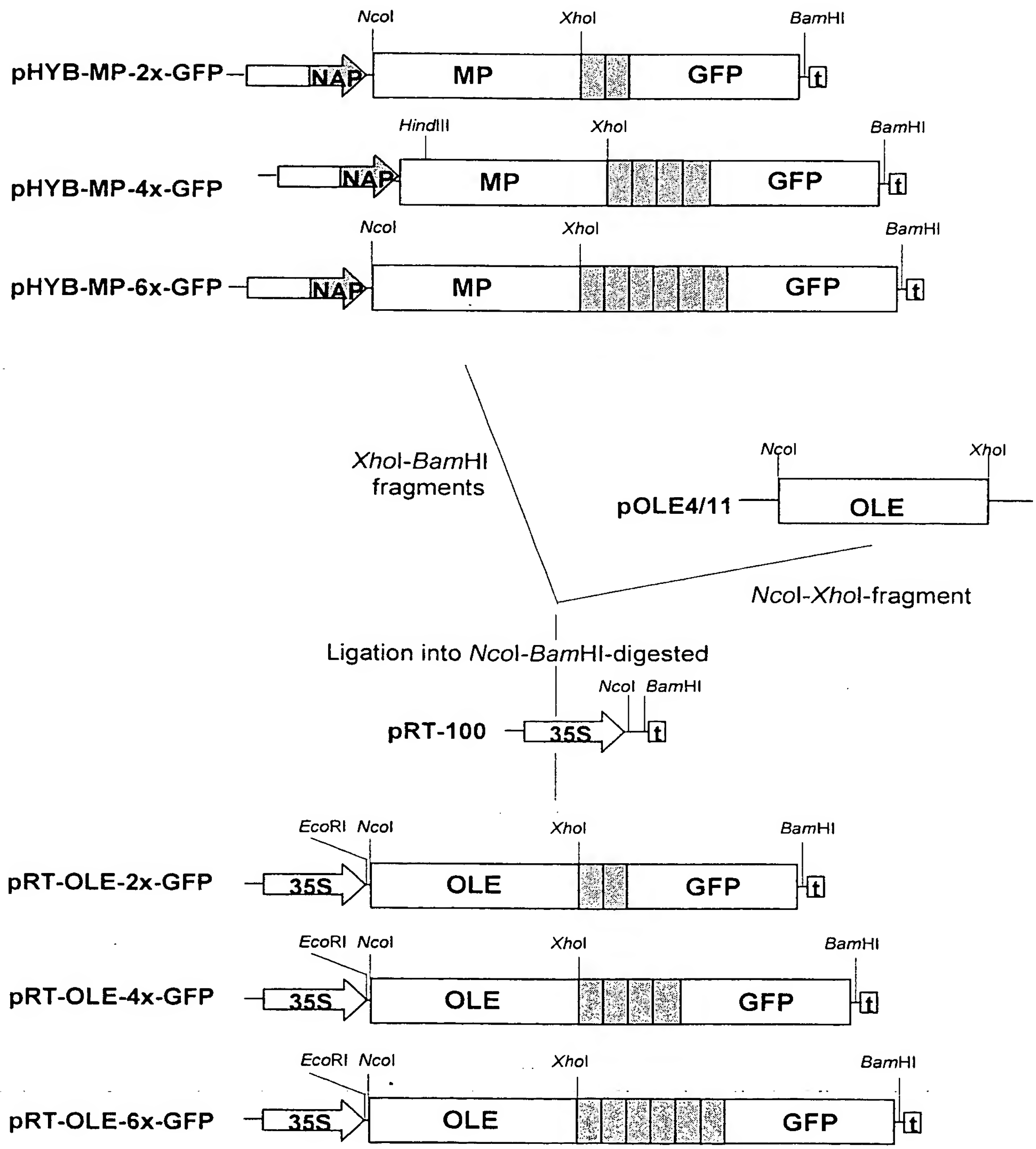
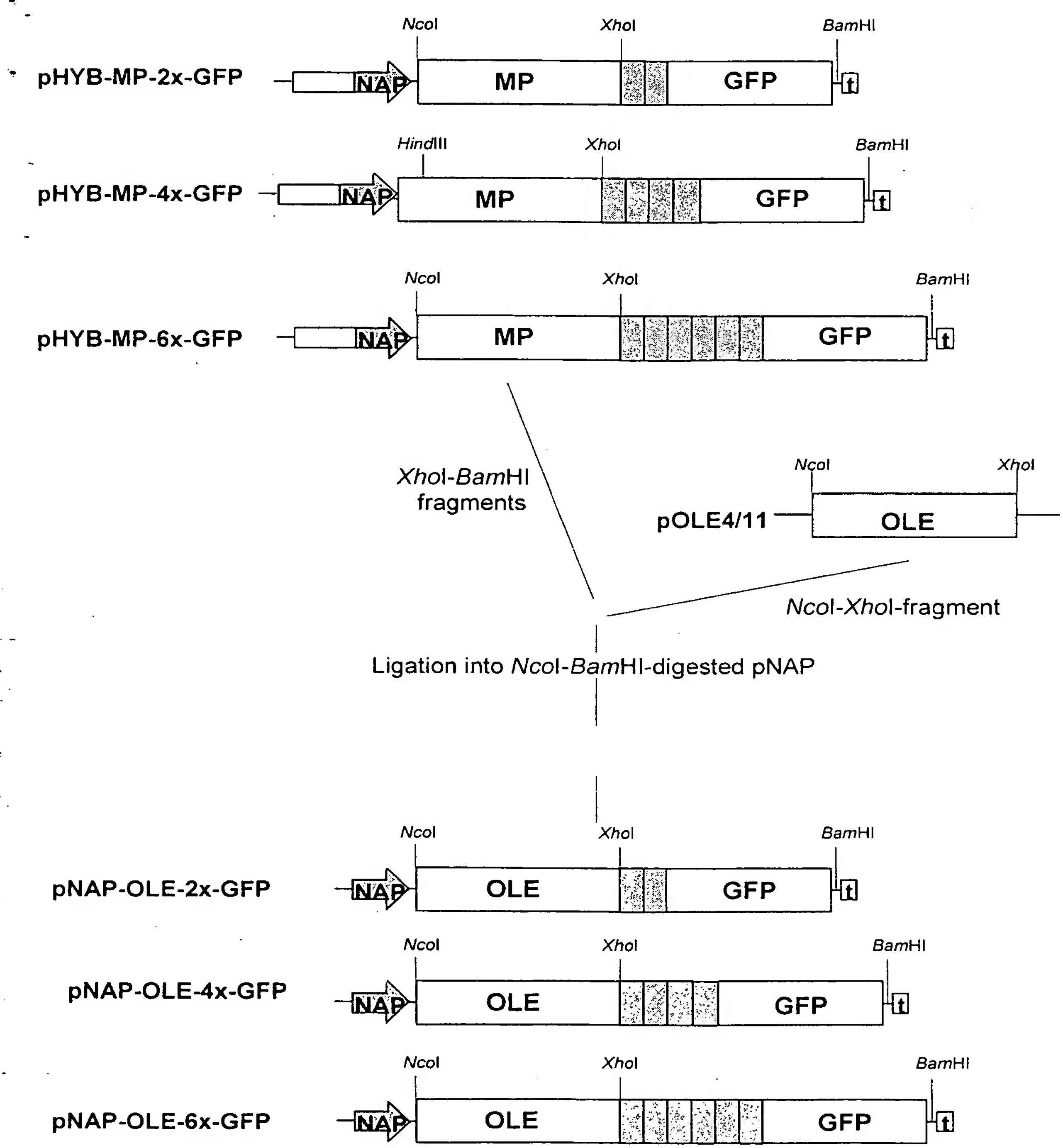

HindIII
XhoI
BamHI
pRT-MP-2x-GFP
35S
MP
GFP
t
pRT-MP-4x-GFP
pRT-MP-6x-GFP
HindIII-BamHI fragments
NcoI
pGEM-30K
NcoI-HindIII-fragment
Ligation into NcoI-BamHI-digested pHYB
pHYB
NAP
pHYB-MP-2x-GFP
pHYB-MP-4x-GFP
pHYB-MP-6x-GFP

Fig. 10


NcoI
XhoI
BamHI
pHYB-MP-2x-GFP
NAP
MP
GFP
t
HindIII
XhoI
BamHI
pHYB-MP-4x-GFP
NAP
MP
GFP
t
NcoI
XhoI
BamHI
pHYB-MP-6x-GFP
NAP
MP
GFP
t
XhoI-BamHI fragments
NcoI
XhoI
pOLE4/11
OLE
NcoI-XhoI-fragment
Ligation into NcoI-BamHI-digested
NcoI BamHI
pRT-100
35S
t
EcoRI NcoI
XhoI
BamHI
pRT-OLE-2x-GFP
35S
OLE
GFP
t
EcoRI NcoI
XhoI
BamHI
pRT-OLE-4x-GFP
35S
OLE
GFP
t
EcoRI NcoI
XhoI
BamHI
pRT-OLE-6x-GFP
35S
OLE
GFP
t

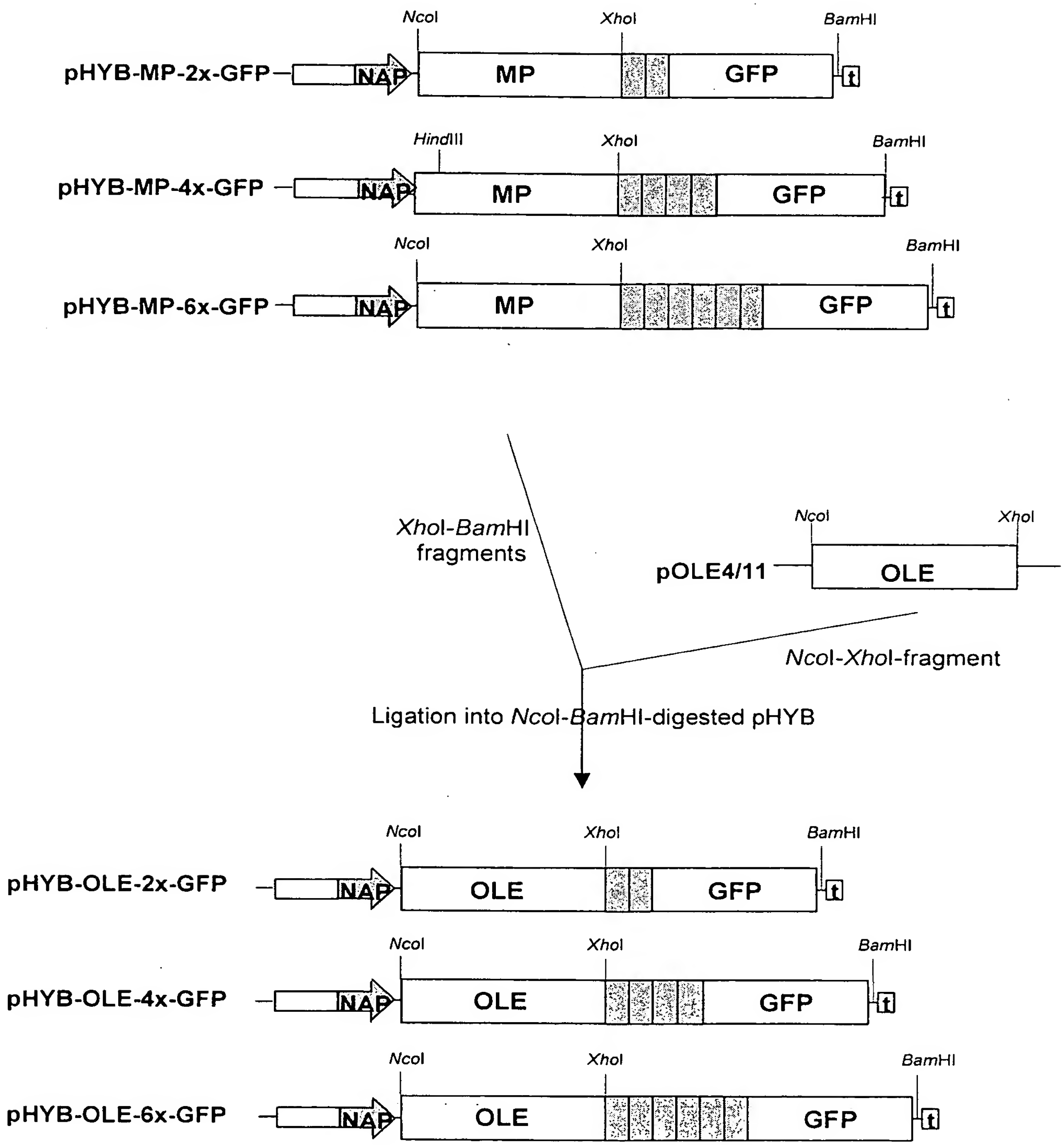
L7

Fig. 11


NcoI
XhoI
BamHI
pHYB-MP-2x-GFP
NAP
MP
GFP
t
HindIII
XhoI
BamHI
pHYB-MP-4x-GFP
NAP
MP
GFP
t
NcoI
XhoI
BamHI
pHYB-MP-6x-GFP
NAP
MP
GFP
t
XhoI-BamHI fragments
NcoI
XhoI
pOLE4/11
OLE
NcoI-XhoI-fragment
Ligation into NcoI-BamHI-digested pNAP
NcoI
XhoI
BamHI
pNAP-OLE-2x-GFP
NAP
OLE
GFP
t
NcoI
XhoI
BamHI
pNAP-OLE-4x-GFP
NAP
OLE
GFP
t
NcoI
XhoI
BamHI
pNAP-OLE-6x-GFP
NAP
OLE
GFP
t

17

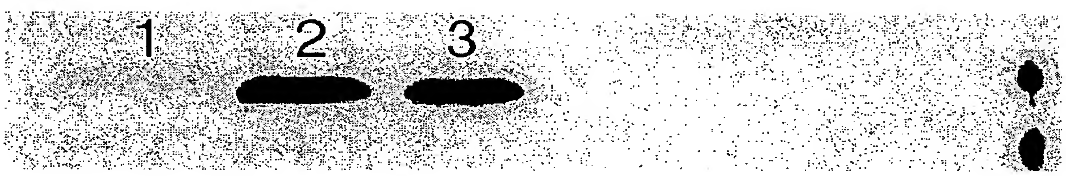
Fig. 12



pHYB-MP-2x-GFP
NcoI
XhoI
BamHI
NAP
MP
GFP
t
pHYB-MP-4x-GFP
HindIII
XhoI
BamHI
NAP
MP
GFP
t
pHYB-MP-6x-GFP
NcoI
XhoI
BamHI
NAP
MP
GFP
t
XhoI-BamHI fragments
pOLE4/11
NcoI
XhoI
OLE
NcoI-XhoI-fragment
Ligation into NcoI-BamHI-digested pHYB
pHYB-OLE-2x-GFP
NcoI
XhoI
BamHI
NAP
OLE
GFP
t
pHYB-OLE-4x-GFP
NcoI
XhoI
BamHI
NAP
OLE
GFP
t
pHYB-OLE-6x-GFP
NcoI
XhoI
BamHI
NAP
OLE
GFP
t

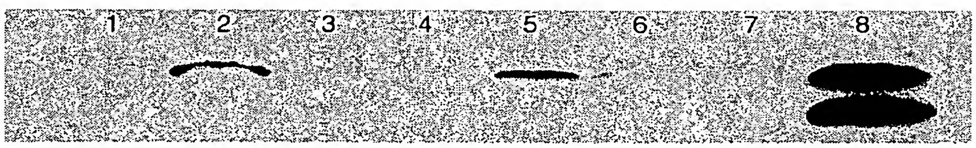
L7

Fig. 13



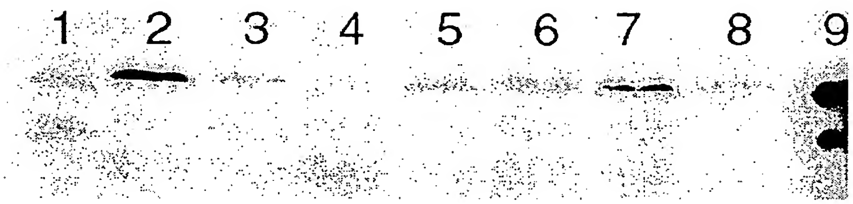
L7

Fig. 14



L7

Fig. 15



L8

SEQUENCE LISTING

<110> Wahlroos, Tony
Atabekov, Joseph
Dorokhov, Yurii
Susi, Petri
Mäkelä, Mauri
Korpela, Timo

<120> Methods and Constructs for Increasing the Content of Selected Amino Acids in Seeds

<130> A2436PFI

<140>
<141>

<160> 17

<170> PatentIn Ver. 2.1

<210> 1
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> H-P-1

<400> 1
gcgcctcgag ttcaccatca ccatcaccat cacgggcacc atcac 45

<210> 2
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> H-P-2

<400> 2

catcaccatc accatgg 17

<210> 3
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> H-M

<400> 3
ccggatccta aagtcgacca tggtgatggt gatggtgatg gtgcc 45

<210> 4
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> C-M-P-1

<400> 4
cacctcgagt atgttgttgc atgtgcatgt gctgttgcat gtcgacaaac 50

<210> 5
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> C-M-P-2

<400> 5
gtttgtcgac atgcaacagc acatgcacat gcaacaacat actcgaggtg 50

<210> 6
<211> 57
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> GL-P-1

<400> 6
gcgcctcgag ttggtggagg tggaggcggt ggaggtggcg tcgacaaatg gatcccc 57

<210> 7
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> GL-P-2

<400> 7
ggggatccat ttgtagacgc cacctcctcc accgcctcca cctccaccaa ctcgaggcgc 60

<210> 8
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> L-P-1

<400> 8
gcgcctcgag ttaaaaagaa aaagaaaaag aaaaagaaaa agaaaaaggt cgacaaatgg 60
atcccc 66

<210> 9
<211> 66

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> L-P-2

<400> 9
ggggatccat ttgtcgacct ttttcttttt ctttttcttt ttctttttct ttttaactcg 60
aggcgc 66

<210> 10
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> OLE-P

<400> 10
aaaaccatgg cggatacagc tagaggaacc catc 34

<210> 11
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> OLE-M

<400> 11
ggggccatgg gagtagtgtg ctggccacca cgagtac 37

<210> 12
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> OLE-3'XhoI

<400> 12
gcgcctcgag aagtagtgtg ctggccacca c 31

<210> 13
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> N30K

<400> 13
gcggaattcc catggctcta gttgttaaag g 31

<210> 14
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> C30K

<400> 14
agacctcgag gaaacgaatc cgattcggcg ac 32

<210> 15
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> OLE-EX-B

<400> 15
tgtgggatcc tacgcaacgg gagagcaccc a 31

<210> 16
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> NAP-P

<400> 16
cttgttagcc atggtttgct atttgtg 27

<210> 17
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic

<220>
<223> NAP-M

<400> 17
tcttactcga gtgaaaccaa attaac 26